АВТОМАТИКА
и
ТЕЛЕМЕХАНИКА

Volume 19, No. 8

August 1958

SOVIET INSTRUMENTATION AND
CONTROL TRANSLATION SERIES

Automation and Remote Control

(The Soviet Journal *Avtomatika i Telemekhanika* in English Translation)

■ This translation of a Soviet journal on automatic control is published as a service to American science and industry. It is sponsored by the Instrument Society of America under a grant in aid from the National Science Foundation, continuing a program initiated by the Massachusetts Institute of Technology.

SOVIET INSTRUMENTATION AND CONTROL TRANSLATION SERIES

■ This translation of the Soviet Journal *Avtomatika i Telemekhanika* is published and distributed at nominal subscription rates under a grant in aid to the Instrument Society of America from the National Science Foundation. This translated journal, and others in the Series (see back cover), will enable American scientists and engineers to be informed of work in the fields of instrumentation, measurement techniques and automatic control reported in the Soviet Union.

The original Russian articles are translated by competent technical personnel. The translations are on a cover-to-cover basis, permitting readers to appraise for themselves the scope, status and importance of the Soviet work.

Publication of *Avtomatika i Telemekhanika* in English translation started under the present auspices in April 1958 with Russian Vol. 18, No. 1 of January 1957. Translation of Vol. 18 has now been completed. The twelve issues of Vol. 19 will be published in English translation by mid-1959.

All views expressed in the translated material are intended to be those of the original authors, and not those of the translators, nor the Instrument Society of America.

Readers are invited to submit communications on the quality of the translations and the content of the articles to ISA headquarters. Pertinent correspondence will be published in the "Letters" section of the ISA Journal. Space will also be made available in the ISA Journal for such replies as may be received from Russian authors to comments or questions by American readers.

Subscription Prices:

Per year (12 issues), starting with Vol. 19, No. 1

General:	*United States and Canada*	$30.00
	Elsewhere	33.00
Libraries of non-profit academic institutions:		
	United States and Canada	$15.00
	Elsewhere	18.00
Single issues to everyone, each		$ 6.00

See back cover for combined subscription to entire Series.

Subscriptions and requests for information on back issues should be addressed to the:

Instrument Society of America
313 Sixth Avenue, Pittsburgh 22, Penna.

Translated and printed by Consultants Bureau, Inc.

Volume XIX No. 8 – August 1958

English Translation Published May 1959

Automation and Remote Control

The Soviet Journal Avtomatika i Telemekhanika in English Translation

Avtomatika i Telemekhanika is a Publication of the Academy of Sciences of the USSR

See following page for Table of Contents.

AUTOMATION AND REMOTE CONTROL

Volume 19, Number 8 August 1958

CONTENTS

FLUCTUATION ACTION IN THE SIMPLEST PARAMETRIC SYSTEMS

V. I. Tikhonov

(Moscow)

The statistical characteristics of the output, random signal of a parametric system are obtained for the case when the external action and the variable parameter are correlated, stationary, normal, random functions. The system under consideration is described by a first-order, linear, differential equation.

In many cases it is not possible to rigorously solve the problem of the action of random signals in nonlinear systems, necessitating, as a consequence, the introduction of certain simplifications. At the same time it is sometimes possible to reduce the problem to the solution of a first-order, linear, differential equation, with a variable coefficient of the form

$$\frac{dy}{dt} + \xi(t)\,y = \eta(t), \tag{1}$$

where $\xi(t)$ and $\eta(t)$ are random functions. Henceforth, we shall assume that these functions are correlated, stationary and normal, with known characteristics.

Let us introduce the following notations:

$$m_1 = \langle \xi(t) \rangle, \quad m_2 = \langle \eta(t) \rangle, \quad \sigma_1\sigma_2 R_{12}(\tau) = \langle [\xi(t) - m_1]\,[\eta(t+\tau) - m_2] \rangle,$$
$$\sigma_1^2 R_1(\tau) = \langle [\xi(t) - m_1]\,[\xi(t+\tau) - m_1] \rangle,$$
$$\sigma_2^2 R_2(\tau) = \langle [\eta(t) - m_2]\,[\eta(t+\tau) - m_2] \rangle,$$

where the skew brackets denote the mathematical expectation operation, i.e., the operation of averaging over many different realizations.

As is well known [1], the general solution of Equation (1) can be expressed in the form of the sum of the general solution y_1 of a homogeneous equation and the particular solution y_2 of an inhomogeneous equation. For the initial conditions $y = y_0$, $t = t_0$, we can write

$$y(t) = y_1(t) + y_2(t) = y_0 \exp\Big(-\int_{t_0}^{t} \xi(x_1)\,dx_1\Big) + \int_{t_0}^{t} \eta(u) \exp\Big(-\int_{u}^{t} \xi(x_2)\,dx_2\Big)\,du. \tag{2}$$

Let us note that this solution of Equation (1) can be obtained in a different way. The solution of the differential equation

$$\frac{dz}{dt} + \xi(t)\,z = \exp[-\alpha\eta(t)],$$

where α is some parameter, for the initial conditions $z = y_0$, $t = t_0$ has the form:

$$z(t) = z_1(t) + z_2(t) = y_0 \exp\Big(-\int_{t_0}^{t} \xi(x_1)\,dx_1\Big) + \int_{t_0}^{t} \exp\Big[-\alpha\eta(u) - \int_{u}^{t} \xi(x_2)\,dx_2\Big]\,du. \tag{3}$$

From a comparison of Expressions (2) and (3), it follows that

$$y(t) = z_1(t) - \left[\frac{\partial}{\partial\alpha} z_2(t)\right]_{\alpha=0}. \tag{4}$$

This result permits a significant simplification in the calculation of statistical characteristics y(t). In particular, for the calculation of various moments it is sufficient to utilize the expression for the n-dimensional characteristic function of a normal random process:

$$\vartheta_n(\Omega_1, \ldots, \Omega_n) = \exp\left[j \sum_{\mu=1}^{n} m_\mu \Omega_\mu - \frac{1}{2} \sum_{\mu,\nu=1}^{n} \sigma_\mu \sigma_\nu R_{\mu\nu}(\tau)\, \Omega_\mu \Omega_\nu\right]. \tag{5}$$

Here $m_\mu = m(t_\mu)$, $\sigma_\mu = \sigma(t_\mu)$ are, respectively, the mean and mean-square values of the random process at some instant of time t_μ, $R_{\mu\nu}(\tau) = R(t_\mu - t_\nu)$ is the coefficient of correlation between two values of the random process at times t_μ and t_ν.

Let us calculate the mean value and the correlation function of y(t). From Relationship (4) we have

$$m(t) = \langle y(t)\rangle = y_0 \left\langle \exp\left(-\int_{t_0}^{t} \xi(x_1)\, dx_1\right)\right\rangle - \left[\frac{\partial}{\partial\alpha}\langle z_2(t)\rangle\right]_{\alpha=0}. \tag{6}$$

Insofar as the linear transformations of a normal process yield a normal process too, then the first term of Equation (6) represents the value of the one-dimensional characteristic function of the normal process

$$\zeta(t) = \int_{t_0}^{t} \xi(x)\, dx, \quad m_1(t - t_0) = \langle \zeta(t)\rangle, \quad \sigma_\zeta^2(t, t_0) = \sigma_1^2 \int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2 - x_1)\, dx_1\, dx_2$$

at the point $\Omega_1 = j$, i.e.,

$$\vartheta_1(j) = \left\langle \exp\left(-\int_{t_0}^{t} \xi(x_1)\, dx_1\right)\right\rangle =$$

$$= \exp\left[-m_1(t - t_0) + \frac{1}{2}\sigma_1^2 \int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2 - x_1)\, dx_1 dx_2\right]. \tag{7}$$

Analogously, making use of Expression (5), we obtain

$$\vartheta_2(j\alpha, j) = \left\langle \exp\left[-\alpha\eta(u) - \int_{u}^{t} \xi(x_2)\, dx_2\right]\right\rangle = \exp\Big\{-[\alpha m_2 + m_1(t-u)] +$$

$$+ \frac{1}{2}\left[\alpha^2\sigma_2^2 + 2\alpha\sigma_2\sigma_\zeta(t, u)\int_{u}^{t} R_{12}(u - x_2)\, dx_2 + \sigma_\zeta^2(t, u)\right]\Big\}.$$

Consequently,

$$\left[\frac{\partial}{\partial\alpha}\langle z_2(t)\rangle\right]_{\alpha=0} = \int_{t_0}^{t}\left[\frac{\partial}{\partial\alpha}\vartheta_2(j\alpha, j)\right]_{\alpha=0} du =$$

$$= \int_{t_0}^{t}\left[-m_2 + \sigma_2\sigma_\zeta(t, u)\int_{u}^{t} R_{12}(u - x_2)\, dx_2\right] \times$$

$$\times \exp\left[-m_1(t - u) + \frac{1}{2}\sigma_1^2 \int_{u}^{t}\int_{u}^{t} R_1(x_2 - x_1)\, dx_1\, dx_2\right]. \tag{8}$$

Substituting (7) and (8) into Equation (6), we obtain

$$m(t) = y_0 \exp\Big[-m_1(t-t_0) + \frac{1}{2}\sigma_1^2 \int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2-x_1)\,dx_1\,dx_2\Big] +$$

$$+ \int_{t_0}^{t}\Big[m_2 - \sigma_1\sigma_2 \int_{t_0}^{t} dx_1 \int_{t_0}^{t} dx_2 \int_{u}^{t} R_1(x_2-x_1)\,R_{12}(u-x_3)\,dx_3\Big] \times$$

$$\times \exp\Big[-m_1(t-u) + \frac{1}{2}\sigma_1^2 \int_{u}^{t}\int_{u}^{t} R_1(x_2-x_1)\,dx_1\,dx_2\Big]\,du. \qquad (9)$$

The correlation function of $y(t)$ is equal to

$$k(t_1, t_2) = M_2(t_1, t_2) - m(t_1)\,m(t_2) = \langle y(t_1)\,y(t_2)\rangle - m(t_1)\,m(t_2), \qquad (10)$$

where $M_2(t_1, t_2)$ is a two-dimensional moment. For it we obtain from Formula (4) the expression

$$M_2(t_1, t_2) = \langle z_1(t_1)\,z_1(t_2)\rangle - \Big[\frac{\partial}{\partial\alpha}\langle z_1(t_1)\,z_2(t_2)\rangle\Big]_{\alpha=0} -$$

$$-\Big[\frac{\partial}{\partial\alpha}\langle z_1(t_2)\,z_2(t_1)\rangle\Big]_{\alpha=0} + \Big\langle\Big[\frac{\partial}{\partial\alpha} z_2(t_1)\Big]_{\alpha=0}\Big[\frac{\partial}{\partial\alpha} z_2(t_2)\Big]_{\alpha=0}\Big\rangle. \qquad (11)$$

In the calculation of the separate summands of the right side, one should make use of the expressions for the three-dimensional and four-dimensional characteristic functions of a normal process, which can be derived from Equation (5) for $n = 3$ and $n = 4$.

Making use of the equations obtained in this manner, it can be shown that

$$\langle z_1(t_1)\,z_1(t_2)\rangle = y_0^2 \exp\Big\{-m_1(t_1+t_2-2t_0) + \frac{1}{2}\Big[\sigma_1^2 \int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2-x_1)\,dx_1\,dx_2 +$$

$$+ 2\sigma_\zeta(t_1, t_0)\,\sigma_\zeta(t_2, t_0) \int_{t_0}^{t_1}\int_{t_0}^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2 + \sigma_1^2 \int_{t_0}^{t_2}\int_{t_0}^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2\Big]\Big\},$$

$$\langle z_1(t_1)\,z_2(t_2)\rangle = y_0 \int_{t_0}^{t} \exp\Big\{-\alpha m_2 - m_1(t_1+t_2-t_0-u) +$$

$$+ \frac{1}{2}\Big[\alpha^2\sigma_2^2 + \sigma_1^2 \int_{t_0}^{t_1}\int_{t_0}^{t_1} R_1(x_2-x_1)\,dx_1\,dx_2 + \sigma_1^2 \int_{u}^{t_2}\int_{u}^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2 + \qquad (12)$$

$$+ 2\alpha\sigma_2\sigma_\zeta(t_1, t_0) \int_{t_0}^{t_1} R_{12}(u-x_1)\,dx_1 + 2\alpha\sigma_2\sigma_\zeta(t_2, u) \int_{u}^{t_2} R_{12}(u-x_2)\,dx_2 +$$

$$+ 2\sigma_\zeta(t_1, t_0)\,\sigma_\zeta(t_2, u) \int_{t_0}^{t}\int_{u}^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2\Big]\Big\}\,du.$$

From here

$$\Big[\frac{\partial}{\partial\alpha}\langle z_1(t_1)\,z_2(t_2)\rangle\Big]_{\alpha=0} = y_0 \int_{t_0}^{t_2}\Big[-m_2 + \sigma_2\sigma_\zeta(t_1, t_0) \int_{t_0}^{t_1} R_{12}(u-x_1)\,dx_1 +$$

$$+ \sigma_2\sigma_\zeta(t_2, u) \int_{u}^{t_2} R_{12}(u-x_2)\,dx_2\Big] \exp\Big\{-m_1(t_1+t_2-t_0-u) + \frac{1}{2}\sigma_\zeta^2(t_1, t_0) +$$

$$+ \frac{1}{2}\sigma_\zeta^2(t_2, u) + \sigma_\zeta(t_1, t_0)\,\sigma_\zeta(t_2, u) \int_{t_0}^{t_1}\int_{u}^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2\Big\}\,du. \qquad (13)$$

Interchanging t_1 and t_2 at this point we obtain the third term of the right side of Equation (11). The last (fourth) term in this expression can be written in the following form which is more convenient for calculation:

$$\left\langle \left[\frac{\partial}{\partial \alpha} z_2(t_1)\right]_{\alpha=0} \left[\frac{\partial}{\partial \alpha} z_2(t_2)\right]_{\alpha=0} \right\rangle = \left[\frac{\partial^2}{\partial \alpha_1 \partial \alpha_2} \langle z_2(t_1) z_2(t_2)\rangle\right]_{\alpha_1=\alpha_2=0}.$$

Applying Equation (5) with n = 4 and carrying out the necessary computations, we obtain

$$\left[\frac{\partial^2}{\partial \alpha_1 \partial \alpha_2} \langle z_2(t_1) z_2(t_2)\rangle\right]_{\alpha_1=\alpha_2=0} = \int_{t_0}^{t_1}\int_{t_0}^{t_2} \Big\{ \sigma_2^2 R_2(u-v) +$$

$$+ \left[m_2 - \sigma_2\sigma_\zeta(t_1, u)\int_u^{t_1} R_{12}(u-x_1)\,dx_1 - \sigma_2\sigma_\zeta(t_2, u)\int_v^{t_2} R_{12}(u-x_2)\,dx_2\right] \times$$

$$\times \left[m_2 - \sigma_2\sigma_\zeta(t_1, v)\int_u^{t_1} R_{12}(v-x_1)\,dx_1 - \sigma_2\sigma_\zeta(t_2, v)\int_v^{t_2} R_{12}(v-x_2)\,dx_2\right]\Big\} \times$$

$$\times \exp\Big\{-m_1(t_1+t_2-u-v) + \frac{1}{2}\sigma_\zeta^2(t_1, u) + \frac{1}{2}\sigma_\zeta^2(t_2, v) +$$

$$+ \sigma_\zeta(t_1, u)\,\sigma_\zeta(t_2, v)\int_u^{t_1}\int_v^{t_2} R_1(x_2-x_1)\,dx_1\,dx_2\Big\}\,du\,dv. \tag{14}$$

Substituting the obtained Expressions (12) - (14) into Equation (11) and then into (10), we obtain a general formula for the correlation function. Since this equation turns out to be very cumbersome, we will limit ourselves here to examining certain particular cases.

1. Assume zero initial conditions ($y_0 = 0$, $t_0 = 0$) and $\xi(t) = m_1$. Then Equation (1) becomes a first-order, linear, differential equation with a random external force. In this case we have from Equation (9)

$$m(t) = m_2\int_0^t e^{-m_1(t-u)}\,du = \frac{m_2}{m_1}(1-e^{-m_1 t}). \tag{15}$$

The mean value of m in a stationary condition ($t \to \infty$) will be finite and equal to m_2/m_1 only when $m_1 > 0$, i.e., if the system possesses positive damping. We can, therefore, say that the system being considered is stable relative to its mean value only when $m_1 > 0$.

For a two-dimensional moment we obtain from Equation (14)

$$M_2(t_1, t_2) = \int_0^{t_1}\int_0^{t_2} [\sigma_2^2 R_2(u-v) + m_2^2]\exp[-m_1(t_1+t_2-u-v)]\,du\,dv.$$

Making use of Equations (10) and (15), it is easy to see that the correlation function is equal to

$$k(t_1, t_2) = \sigma_2^2\int_0^{t_1}\int_0^{t_2} R_2(u-v)\exp[-m_1(t_1+t_2-u-v)]\,du\,dv.$$

Introducing new variables $x = t_1 - u$, $y = t_2 - v$, $\tau = t_1 - t_2$, $t = t_1$, we can write

$$k(t, t-\tau) = \sigma_2^2\int_0^{t}\int_0^{t-\tau} R_2(\tau+y-x)\exp[-m_1(x+y)]\,dx\,dy.$$

Let $R_2(\tau + y - x) = \exp(-\beta\,|\tau + y - x|)$, for example. Carrying out separate integrations for $\tau > 0$ (Fig. 1, a) and $\tau < 0$ (Fig. 1, b), by splitting up the regions of integration, in each case into two sub-regions, depending on the sign of the argument of the coefficient of correlation, we obtain

$$k(t,\, t-\tau) = \begin{cases} \dfrac{\sigma_2^2}{m_1^2-\beta^2}\Big\{e^{m_1\tau}\left(e^{-2m_1t} - e^{-(m_1+\beta)t}\right) + \\ \quad + \dfrac{\beta}{m_1} e^{m_1\tau}\left(e^{-2m_1t}-1\right) - e^{\beta\tau}\left(e^{-(m_1+\beta)t}-1\right)\Big\},\ \tau < 0 \\ \dfrac{\sigma_2^2}{m_1^2-\beta^2}\Big\{e^{-2m_1t+m_1\tau} - e^{-(m_1+\beta)t+\beta\tau} - e^{-(m_1+\beta)t+m_1\tau} + e^{-\beta\tau} + \\ \quad + \dfrac{\beta}{m_1} e^{-m_1\tau}\left(e^{-2m_1(t-\tau)}-1\right)\Big\},\ \tau > 0. \end{cases}$$

From this equation it follows that when $m_1 > 0$ there is a stationary value of the correlation function

$$k(\tau) = \frac{\sigma_2^2}{m_1^2-\beta^2}\left(e^{-\beta|\tau|} - \frac{\beta}{m_1} e^{-m_1|\tau|}\right).$$

2. Let us suppose $\eta(t) \equiv 0$. In this case the mean value will be, according to Equation (9)

$$m(t) = y_0 \exp\left[-m_1(t-t_0) + \frac{1}{2}\sigma_1^2 \int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2-x_1)\,dx_1\,dx_2\right].$$

Substituting in the variables $x_2 - x_1 = s$, $x_1 = x$, and carrying out the integration, we obtain

$$m(t) = y_0 \exp\left[-m_1(t-t_0) + \sigma_1^2 \int_0^{t-t_0} (t-t_0-s)\,R_1(s)\,ds\right].$$

To determine the stationary condition we must let $(t - t_0) \to \infty$. At the same time, it is apparent from the equation that the mean value for the stationary condition is equal to zero only if

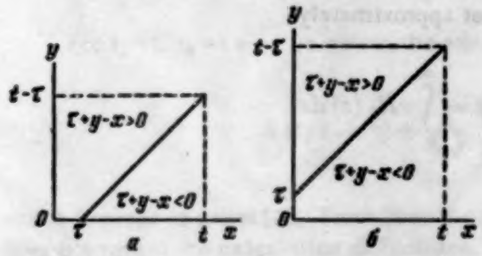


Fig. 1

$$m_1 - \sigma_1^2 \int_0^{\infty} R_1(s)\,dt > 0. \qquad (16)$$

For the reverse inequality the mean value tends to infinity and the system becomes unstable in this sense.

Making use of Equation (12) it is easy to see that, for the given conditions, the correlation function is equal to

$$k(t_1,\, t_2) = y_0^2 m(t_1) m(t_2) \left\{\exp\left[2\sigma_\zeta(t_1,\, t_0)\,\sigma_\zeta(t_2,\, t_0) \int_{t_0}^{t_2}\int_{t_0}^{t_1} R_1(x_2-x_1)\,dx_1\,dx_2\right] - 1\right\}.$$

From this we obtain the dispersion

$$\sigma^2(t) = y_0^2 m^2(t) \left\{\exp\left[2\sigma_1^2\left(\int_{t_0}^{t}\int_{t_0}^{t} R_1(x_2-x_1)\,dx_1\,dx_2\right)^2\right] - 1\right\} =$$

$$= y_0^2 m^2(t) \left\{\exp\left[8\sigma_1^2\left(\int_0^{t-t_0} (t-t_0-s)\,R_1(s)\,ds\right)^2\right] - 1\right\}.$$

Insofar as for finite values of m_1 and σ_1, for sufficiently large $t > 0$, the inequality

$$\left[m_1 - \sigma_1^2 \int_0^\infty R_1(s)\,ds\right](t-t_0) < \sigma_1^2\left(\int_0^\infty R_1(s)\,ds\right)^2 (t-t_0)^2,$$

is always satisfied then, beginning at some instant of time, the dispersion will grow without limit and the system being examined (for zero initial conditions) turns out to be unstable with respect to the dispersion, despite the fact that when Condition (16) is satisfied the mean value tends to zero.

There is a physical explanation for such a somewhat unusual result. Since the fluctuations $\xi(t)$ are assumed to be normal, with mean value m_1 and dispersion σ_1^2, then there exist random intervals of time Δ_i' (Fig. 2) during which $\xi(t) > 0$, and $y(t)$, subject to the initial deviation y_0, dies away. However, besides the intervals Δ_i' there are intervals Δ_k'' during which $\xi(t) \leq 0$, and, therefore, during them the deviation $y(t)$ grows. Despite the fact that the general time $\Sigma\Delta_i'$, during which $\xi(t) > 0$, is greater than the general time $\Sigma\Delta_k''$, during which $\xi(t) \leq 0$, the dispersion of $y(t)$ grows with time as a result of the fact that for $\xi(t) \approx 0$ the rate of growth of $y(t)$ is very large.

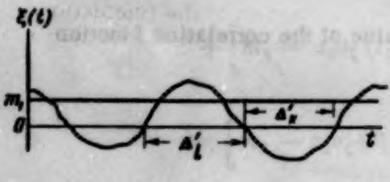


Fig. 2

3. Let the initial conditions be zero ($y_0 = 0$, $t_0 = 0$) and the random functions $\xi(t)$ and $\eta(t)$ be noncorrelated. Then from Equation (9), we have

$$m(t) = m_2 \int_0^t \exp\left[-m_1(t-u) + \frac{1}{2}\sigma_1^2 \int_u^t\int_u^t R_1(x_2 - x_1)\,dx_1\,dx_2\right] du =$$

$$= m_2 \int_0^t \exp\left[-m_1 v + \sigma_1^2 \int_0^v (v-s)\,R_1(s)\,ds\right] dv. \tag{17}$$

To determine the condition for the existence of a finite, mean value in the stationary condition ($t \to \infty$) we divide the interval of integration with respect to v into two sub-intervals. Let us pick some value of T which is finite but sufficiently large for $TR_1(T) \approx 0$. Then we can set approximately

$$\int_0^T R_1(s)\,ds = \int_0^\infty R_1(s)\,ds, \qquad \int_0^T sR_1(s)\,ds = \int_0^\infty sR_1(s)\,ds.$$

We can write

$$m(t) = m_2 \int_0^T \exp\left[-m_1 v + \sigma_1^2 \int_0^v (v-s)\,R_1(s)\,ds\right] dv +$$

$$+ m_2 \int_T^\infty \exp\left\{-\left(m_1 - \sigma_1^2 \int_0^\infty R_1(s)\,ds\right)v - \int_0^\infty sR_1(s)\,ds\right\} dv.$$

From here it follows that the mean value for the stationary condition will be finite if Condition (16) is satisfied.

Letting the coefficient of mutual correlation $R_{12}(\tau) = 0$ in Equation (14), we obtain the equation for the two-dimensional moment

$$M_2(t_1, t_2) = \int_0^{t_1}\int_0^{t_2} [\sigma_2^2 R_2(u-v) + m_2^2] \exp\left\{-m_1(t_1 + t_2 - u - v) + \right.$$

$$+ \frac{1}{2}\sigma_\zeta^2(t_1, u) + \frac{1}{2}\sigma_\zeta^2(t_2, v) +$$

$$\left. + \sigma_\zeta(t_1, u)\,\sigma_\zeta(t_2, v) \int_u^{t_1}\int_v^{t_2} R_1(x_2 - x_1)\,dx_1\,dx_2\right\} du\,dv. \tag{18}$$

Without giving the coefficients of correlation R_1 and R_2 concrete values it is difficult to find the conditions for which, for example, the dispersion in the stationary condition has a finite value.

It should be noted that particular cases of the example in question were analyzed in papers [2, 3].

In paper [2] the one-dimensional probability density for the process y(t) was derived for zero initial conditions, with y(t) determined by the differential equation $dy/dt + \xi(t)y = \xi(t)h_0$, where h_0 is a unit step function.

The results of paper [3] are applicable to the example under consideration if $m_2 = 0$. Let us show this.

It follows from Expression (17) that in this case $m(t) = 0$ and Equation (18) gives the correlation function which can be written in another form. Let us introduce instead of the correlation function of the fluctuations $\eta(t)$ their spectral intensity $S_2(\omega)$ by means of the well known formula

$$\sigma_2^2 R_2(u-v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_2(\omega)\, e^{j\omega(u-v)}\, d\omega. \qquad (19)$$

Let $H(j\omega, t)e^{j\omega t}$ be the response of System (1) to the harmonic signal $e^{j\omega t}$. For zero initial conditions we have from Expression (2)

$$H(j\omega, t) = e^{-j\omega t}\int_0^t \exp\left(j\omega u - \int_u^t \xi(x)\,dx\right) du.$$

It is not difficult to verify that, after substituting Expression (19), the exponential multiplier appearing under the integral sign in Equation (18) can be represented in the form of the mean value of the product $H(j\omega, t_1)H(-j\omega, t_2)$. Then Equation (18) can be written in the form

$$k(t_1, t_2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_2(\omega)\,\langle H(j\omega, t_1)\,H(-j\omega, t_2)\rangle\, e^{-j\omega(t_2-t_1)}\, d\omega.$$

For $t_1 = t$, $t_2 = t + \tau$, we obtain the equation

$$k(t, t+\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_2(\omega)\,\langle H(j\omega, t)\,H(-j\omega, t+\tau)\rangle\, e^{-j\omega\tau}\, d\omega,$$

which is given in paper [3]. From here it can be seen that the last equation is of a purely formal nature and does not reveal the calculation difficulties.

In concluding we will note that the formulas obtained above can be generalized for the case when the correlated, random functions $\xi(t)$ and $\eta(t)$ are not normal but possess a probability density in the form of many dimensional Edgeworth series [4]. However, the final formulas and calculations turn out, in this case, to be very cumbersome and unwieldy and are therefore not given here.

LITERATURE CITED

[1] N. M. Matveev, in the book: Methods of Integrating Ordinary Differential Equations [In Russian] (Leningrad University Press, 1955).

[2] A. Rosenblom, T. Heilfron, and D. L. Trautman, "Analysis of linear systems with randomly varying inputs and parameters," IRE Convention record, 4, 106 (1954).

[3] L. A. Zadeh, "Correlation functions and power spectra in variable networks," Proc. IRE 38, 11, patt. 1342 (1950).

[4] P. I. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, "Quasimomentary functions in the theory of random processes," [In Russian] DAN SSSR 94, 4 (1954).

Received October 28, 1957

A METHOD OF SOLVING A PARTICULAR CLASS OF INTEGRAL EQUATIONS BY CASING COMPUTERS

Iu. S. Val'denberg

(Moscow)

Zeidel's iterative method of the approximate solution of Friedholm's and Volterra's first and second kind of linear integral equations is presented. The feasibility of using computers for the solution of special types of these equations is indicated. The application of this method to some problems of synthesis of automatic control systems is examined.

INTRODUCTION

In the solution of problems of synthesis of automatic control systems (ACS), to find their optimum characteristics it is necessary to solve integral equations [1]. The necessity of mechanizing the solution of these equations stimulated the development of a method for their solution using computers. The essential nature of such a project is also determined by the fact that a large number of problems in many fields of science and technology leads to integral equations. In the present work a method is given for the solution, by means of a computer, of Friedholm's and Volterra's first and second kind of linear integral equations, having a convolution-type kernel

$$\varphi(x) = f(x) - \lambda \int_0^{b \text{ or } x} K(x-s)\varphi(s)\,ds, \tag{1}$$

$$f(x) = \int_0^{b \text{ or } x} K(x-s)\varphi(s)\,ds, \tag{2}$$

where the given functions are the kernel $K(x-s)$ and $f(x)$ the right side of the equation, $\varphi(x)$ is the solution, $0 \leq x, s \leq b$. The given functions are relationships obtained experimentally and introduced into the machine without a preliminary approximation. The problem of finding a unique solution is solved.

The integral equations are solved approximately by the method of successive approximations, i.e., the iterative method [2, 4].

Method of Solving Equations

The iterative method of solving integral equations is used not only in numerical solutions but also in special computers. This method is very convenient for use in computers because of the uniformity of the mathematical algorism.

Two basic types of the iterative process are known: ordinary iteration and Zeidel's iteration (or Gauss-Zeidel) [2, 3]. In the case of Zeidel's iteration, when the m-th approximation for the i-th ordinate of the sought after function is calculated, the earlier found 1-st, 2-nd, . . . , (i − 1)-th ordinates of the m-th approximation for the sought after function are taken into account; while the i-th, (i + 1)-th, etc. ordinates are taken from the

preceding (m − 1)-th approximation. In other words, refinements of the results of the calculations are carried out after each ordinate of the sought after function is determined, and not after the calculation of the given approximation for the whole sought after function, as is done in ordinary iteration. This explains why Zeidel's methods leads, in many cases, to a much more rapid convergence than the one arising from the ordinary method of iteration.

Up until now, the method of ordinary iteration [4, 6, 8, 9] has been used in computers developed for solving integral equations. Zeidel's iterative method has been given preference in this paper. Besides speeding up the convergence, as mentioned above, this type of iteration permits one to construct an iterative process in a form which secures the convergence of Friedholm's first kind of equation with a symmetrical kernel, the quadratic form of which is positively determined. In the case of ordinary iteration convergence of the process may not be guaranteed for these equations. Besides this, Zeidel's iteration, as will be shown later, leads to simpler computer circuits than those arising in ordinary iteration.

The given integral equation is approximated by a system consisting of 32 linear algebraic equations. This system is solved by Zeidel's iterative process. The suitability of the approximation for equations of the 2-nd kind is obvious, while for equations of the first kind it is valid in the case of symmetrical kernels, the quadratic form of which is positively determined. It is easy to show that for such kernels the solution of the integral equation is unique (if it exists), which permits one to approximate the integral by replacing it by a finite summation.

The approximating system is solved by the iterative method using the following algorism [3]:

$$\varphi_i^{(k)} = f_i - 0{,}5\,\Delta s K_{i1}\varphi_1^{(k)} - \Delta s \sum_{j=2}^{i-1} K_{ij}\varphi_j^{(k)} - \Delta s \sum_{j=i}^{n-1} K_{ij}\varphi_j^{(k-1)}, \tag{3}$$

$$\varphi_i^{(k)} = \frac{f_i}{\Delta s K_{ii}} - \frac{1}{K_{ii}}\left[0.5K_{i1}\varphi_1^{(k)} + \sum_{j=2}^{i-1} K_{ij}\varphi_j^{(k)} + \sum_{j=i+1}^{n-1} K_{ij}\varphi_j^{(k-1)}\right], \tag{4}$$

$$\varphi_i^{(k)} = \frac{f_i}{1-\Delta s} - \frac{\Delta s}{1-\Delta s}\left[0.5K_{i1}\varphi_1^{(k)} + \sum_{j=2}^{i-1} K_{ij}\varphi_j^{(k)} + \varphi_i^{(k-1)}\right], \tag{5}$$

$$\varphi_i = \frac{f_i}{\Delta s K_{ii}} - \frac{1}{K_{ii}}\sum_{j=1}^{i-1} K_{ij}\varphi_j \quad \left(i=1,\,2,\,3,\,\ldots,\,n;\ \Delta s = \frac{b}{n-1};\ n=32\right). \tag{6}$$

Here the index in the raised brackets indicates the number of the iterative step, K_{ij} is the matrix approximating the kernel $K(x-s)$. Equation (6) is associated with Volterra's equation of the first kind, and represents a recursive relationship realizing the method of successive substitutions.

Functions arising in integral equations are, as a rule, continuous and have continuous derivatives. In isolated cases where this is not so, special artificial methods have to be used.

A zero vector is used as the initial approximation of the iterative process. In those cases where the zero vector leads to a diverging process (although these cases are possible, they are relatively rare) the integral equation is solved in two stages [4]. First, a rough and very approximate solution of the integral equation is found by the minimizing method (i.e., along an open cycle) in accordance with the following algorisms:

$$\sum_{i=1}^{n}\left[f_i - 0{,}5\,\Delta s K_{i1}\varphi_1 - \Delta s \sum_{j=2}^{n-1} K_{ij}\varphi_j\right]^2 = \varepsilon^2 \to \min, \tag{7}$$

$$\sum_{i=1}^{n}\left[\left|f_i - 0{,}5\Delta s K_{i1}\varphi_1 - \Delta s \sum_{\substack{j=2\\ j\neq i}}^{n-1} K_{ji}\varphi_j - (\Delta s K_{ii} + 1)\,\varphi_i\right|\right] = \delta \to \min. \tag{8}$$

In the second stage a more accurate solution of the integral equation is obtained by the iterative method, with the rough solution, obtained in the first step, being used as the initial approximation. The introduction of an extra step greatly increases the method's possibilities since it permits one to solve, for example, Friedholm's equation of the second kind for cases where a parameter of the equation is greater than the minimum eigenvalue.

As an illustration of the nature of the convergence of the iterative process we give three numerical examples of equations solved in accordance with (3) - (5).

$$\varphi(x) = e^{-x}(x+1.5) - 0.08265\, e^{x} - \int_0^{0.9} e^{-|x-s|}\varphi(s)\,ds,$$

$$\varphi(x) = e^{-x}, \quad n = 30, \quad \varphi_i^{(0)} = 0, \; k = 6,$$

$$\varphi(x) = 8e^{-5x} - 8\int_0^{\infty} e^{-5|x-s|}\varphi(s)\,ds, \quad \varphi(x) = 5.24\, e^{-10.25x}, \; n = 30,$$

$$\varphi_i^{(0)} = 0, \; s_{\max} = 0.6, \; k = 8,$$

$$\varphi(x) = 1 + \int_0^{x} e^{x-s}\varphi(s)\,ds, \quad \varphi(x) = 0.5\,(1 + e^{2x}), \; n = 30, \; \varphi_i^{(0)} = 0, \; k = 5.$$

Here k is the number of iterative steps. The error in the solution of these equations does not exceed 1 %.

Experimental Verification of the Convergence of the Iterative Process

One can find in the literature descriptions of several devices for the solution of integral equations realizing the method of simple iteration [4, 6, 8, 9]. These computers contain, as a rule, two basic circuits: the calculating unit and the memory unit (see, for example, [9]). The calculating unit carries out the multiplication, integration and subtraction operations in accordance with Equations (1) and (2). The memory unit memorizes the ordinates of the sought after function which are determined in the calculator, and which are thereupon reintroduced into the calculating unit. In this way, the iterative process is realized with a closed system consisting of two units – calculating and memory.

In the case of ordinary iteration the memory device consists of two memory units, one of which stores the k-th, while the other stores the (k + 1)-th approximation of the sought after function. After the calculation of each approximation, a special switching device switches these units in such a way that the (k + 2)-th approximation is recorded in the unit in which the k-th approximation was stored [9].

One, and not two, memory units are required for the realization of Zeidel's iterative method and, moreover, the switching circuit is not required.

Any of the known methods of analog computation can be used in the calculating unit of the device. Thus, for example, standard electrical model units were used to this end in [9].

Operational memory devices, found in general purpose digital computers, can be used as the memory unit. The memory cores in such devices usually have two stable conditions, which guarantees the stable operation of the whole device. The location device can be replaced by a simpler arrangement which guarantees the successive scanning of the cores in the memory device.

The following two problems [5] can be considered to be the basic problems which arise in the development of a computer for the solution of integral equations.

1. The problem of designing a memory device. Because Zeidel's iterative method is used, the problem can be solved with relatively simple circuits.

2. The problem of designing the generator of a kernel which is a function of two independent variables. In the general case, for an arbitrary form of kernel, the problem is soluble but leads to very complicated circuitry. For convolution-type kernels it is not necessary to design special equipment, and one can make use of a conventional generator of a function of one argument; while the convolution is achieved by the proper distribution in time of pulses which control the operation of the apparatus. In this case, experimental relationships are used as the given functions in the integral equation. The fact that the problem is soluble only for convolution-type kernels has been completely justified in practice, since many problems of analysis and synthesis of ACS which reduce to integral equations, lead to a convolution-type kernel.

The principles which form the basis of the method (convergence of Zeidel's iterative process, stability of the closed system, etc.) were verified experimentally on a special apparatus, the schematic of which is shown in Fig. 1.

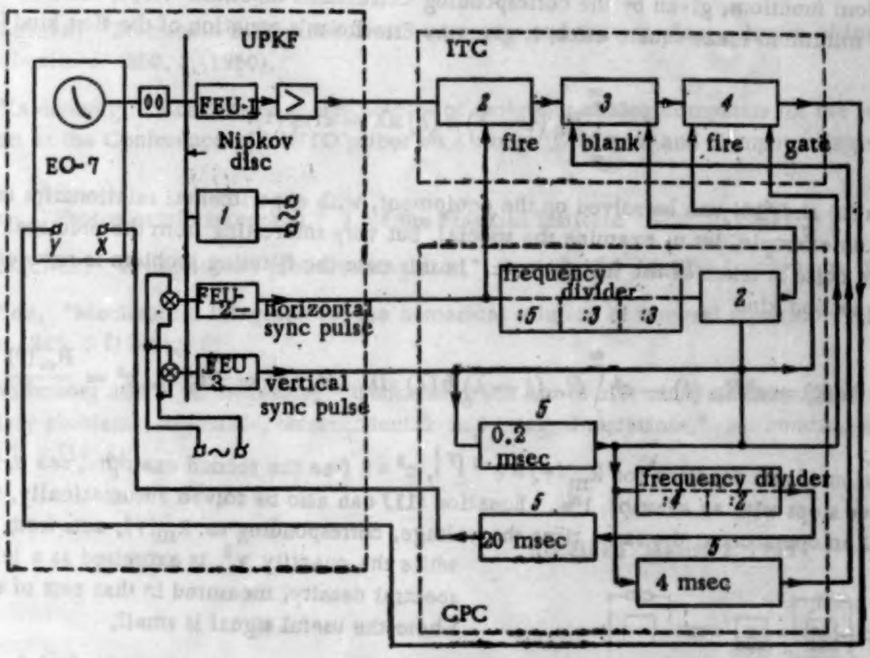


Fig. 1. Circuit of a unit for the verification of the convergence of the iterative process: 1) motor, 2) trigger circuits, 3) integrator, 4) storage device, 5) kipp oscillator.

Here the persistence of a cathode-ray tube is used as the memory unit. The function being recorded is fed to the deflection plates of the tube. The trace of the beam does not disappear at once and the image of the function is read by a photomultiplier FEU-1 with the help of a Nipkov scanning disc with holes distributed in a spiral.

Photomultipliers FEU-2 and FEU-3 are used to obtain the horizontal sync and vertical sync pulses. The integral transformation computer (ITC) integrates the function under the integral sign. The control pulse circuit (CPC) consists of frequency dividers and a kipp oscillator, and produces the pulses which ensure the normal operation of the equipment.

The apparatus was used to solve Volterra's equation of the 2-nd kind $\varphi(x) = 1 - \int_0^x \varphi(x - s)ds$.

The solution $\varphi(x) = e^{-x}$ was obtained in 2,5 sec with an error of 5-6%.

Possible Applications of the Computer

In the synthesis of ACS it is necessary to find the characteristics of a system satisfying the given technical conditions. The transfer function $\Phi(s)$ or the pulse transient function $k(t)$ is such a characteristic. In the solution of this synthesis problem the structure of the system being synthesized is usually not known beforehand and, therefore, the corresponding differential equation cannot be set up.

The characteristic of a linear system which realizes the transformation of definite actions, is given by Volterra's equation of the first kind with a convolution-type kernel

$$x(t) = \int_0^t g(t - \tau) k(\tau) d\tau. \tag{9}$$

The synthesis problem in the indicated statement is very actual for the case when the system is acted on by signals which are random functions of time. In this case the synthesis problem includes the problem of finding the optimum (in a definite sense) characteristics of the system. Thus, for example, if the acting signals are stationary random functions, given by the corresponding correlation functions $R(\tau)$, the $k(t)$ of the system, resulting in the minimum mean square error, is given by Friedholm's equation of the first kind [1]

$$\int_0^\infty R_\varphi(\tau-\lambda)\,k(\lambda)\,d\lambda = R_{h\varphi}(\tau). \tag{10}$$

Equations (9) and (10) can be solved on the equipment, with experimental relationships used as initial conditions. As an example, let us examine the special, but very interesting from the practical point of view, case when white noise is taken as the interference. In this case the filtering problem is reduced to Friedholm's equation of the second kind

$$k(t) = c^2 R_m(t) - c^2\int_0^\infty R_m(t-\lambda)\,k(\lambda)\,d\lambda,\quad R_n(\tau) = \gamma^2\delta(\tau),\quad c^2 = \frac{R_m(0)}{\gamma^2}. \tag{11}$$

The solution of this equation for $R_m(\tau) = e^{-5|\tau|}$, $c^2 = 8$ (see the second example, see p. 714) is obtained in eight iterative steps with an error of 1%. Equation (11) can also be solved automatically, without the participation of an operator; at the same time the voltage, corresponding to $R_m(\tau)$, acts from the correlator, while the quantity γ^2 is expressed as a level of the signal's spectral density, measured in that part of the spectrum where the useful signal is small.

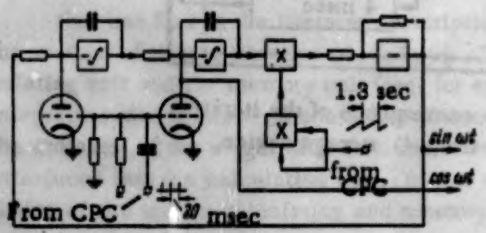


Fig. 2. Circuit for obtaining the kernel $K(x, s) = \cos xs$ and $K(x, s) = \sin xs$.

Methods of solving Equation (10) for other concrete cases may require additional refinement.

Fields of possible application of the equipment are not limited to the solution of integral equations with a convolution-type kernel and with experimental curves (e.g. correlograms). It is also possible to solve integral equations with analytically defined kernels. These kernels are realized on analog computer amplifiers which permit repetitive operations at a frequency of 50 cycles, while the control pulses, determining the work cycle, are formed in the control pulse circuit (CPC). Figure 2 gives an example of the use of such amplifiers for the calculation of Fourier transforms. The kernels $\cos\omega t$ and $\sin\omega t$ are obtained as solutions of the determining differential equation $x'' + k^2x = 0$ [7]. This can be used, for example, when it is required to find the transfer function of a system and the solution of the integral equation gives the pulse transient function. The Fourier transform of $k(t)$ permits one to obtain the transfer function directly.

SUMMARY

1. An apparatus is developed for the solution of a definite class of integral equations by means of Zeidel's iterative process.

2. Provision is made for the solution of integral equations with experimental curves and convolution-type kernels, as well as with analytically established kernels realized on an analog computer amplifier.

In conclusion I must express my thanks to V. V. Solodovnikov for his guidance during the execution of this project.

LITERATURE CITED

[1] V. V. Solodovnikov, in the book: Introduction to the Statistical Dynamics of Automatic Control Systems [In Russian] (Gostekhizdat, 1952).

[2] A. E. Kobrinskii, in the book: Analog Computer Machines [In Russian] (Gostekhizdat, 1954).

[3] V. N. Faddeeva, In the book: Computational Methods of Linear Algebra [In Russian] (Gostekhizdat, 1950).

[4] An. Wallman, "Electronic integraltransform computer and the practical solution of integral equations," J. of the Franklin Institute 250, 1 (1950).

[5] Iu. S. Valdenberg, "Concerning the principles of designing analog computers for the solution of integral equations," Report at the Conference of VNITO pribor on Automatic Control and Computer Engineering [In Russian] 1957.

[6] A. Gray, "Photoelectric integraph," J. of the Franklin Institute 212, 1 (1931).

[7] A. B. MacNee, "A high speed product integrator," Review of Scientific Instruments 24, 3 (1953).

[8] P. A. Tea, "Mechanical integraph for the numerical solution of integral equations," J. of the Franklin Institute 245, 5 (1948).

[9] I. M. Vitenberg and E. A. Gluzberg, "Concerning the use of electrical mathematical models for the solution of boundary problems, algebraic, transcendental and integral equations," Automation and Remote Control (USSR) 17, 7 (1956).

Received March 2, 1957

AUTOMATIC OPTIMALIZER

A. A. Fel'dbaum

(Moscow)

The paper deals with the problem of constructing a machine which minimizes a function of several variables $Q(x_1, \ldots, x_n)$ in the presence of the additional restrictions $H_j(x_1, \ldots, x_n) \le 0$ $(j = 1, \ldots, m)$. Several possible ways of solving this problem are examined, and the circuits of different models of the machine are presented.

1. Statement of the Problem

Self-adjusting systems, which are at present in an early stage of development, will apparently become very important in the next few years.

One of the most important classes of self-adjusting systems are automatic optimization systems, the purpose of which is to find the conditions for a minimum in some quantity Q, usually under additional restrictions. Let us examine the case when Q is a function of the variables $x_1, \ldots, x_n$, and firstly the variables x_i (Fig. 1, a) are set up at the input to the unit O, after which, following a certain elapsed time, the quantity Q appears at the output of the unit:

$$Q = Q(x_1, x_2, \ldots, x_n). \tag{1}$$

The function Q is not known beforehand and the values of x_i, which minimize Q, are unknown.

In Fig. 1 the apparatus A automatically selects the quantities $x_1, \ldots, x_n$ so that a minimum in the quantity Q would be obtained for the general case – with the existence of the additional restrictions, in the form of inequalities, which the variables x_i must satisfy:

$$H_j(x_1, \ldots, x_n) \leqslant 0 \quad (j = 1, \ldots, m). \tag{2}$$

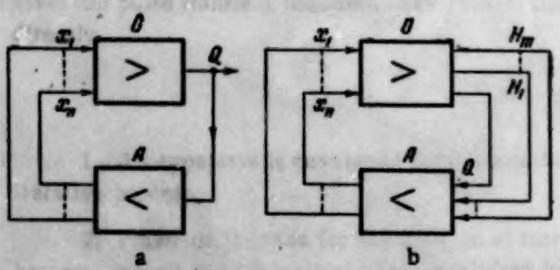


Fig. 1

The number of restrictions m is arbitrary; the three possible cases are m < n, m = n, and m > n. An apparatus which carries out such an operation is subsequently called an automatic optimalizer. In the absence of restrictions the output Q of the unit O is connected to the input of the automatic optimalizer A (Fig. 1, a). If restrictions do exist, the boundary values H_j $(j = 1, \ldots, m)$ are also fed from the output of O to the inputs of the automatic optimalizer A (Fig. 1, b).

Let us assume that Q has only one minimum or that the initial conditions do not originate in a definite region of attraction to the one minimum of Q. In the presence of several minima the described apparatus becomes more complicated; however, it is not difficult to extend the principles being examined to the case of

several minima. Also, let the functions Q and H_j be continuous and differentiable over the whole region of variations of x_i.*

Automatic optimalizers can be utilized as machines for the direct solution of variation problems, for the automatic synthesis of optimum control parts of automatic systems, for the automatic determination of the dynamic characteristics of complicated (for example, nonlinear) members, as well as for the optimization of the behavior of controlled objects under production conditions.

2. Possible Methods of Designing an Automatic Optimalizer

The basic method of solving the given problem is the determination of a method of automatic hunting. In [1, 2] some of the hunting methods were examined: the method of successive changing of the variables, the relaxation method, etc. The following methods are effective for the machine described below.

a. The gradient method, in which the gradient vector is first determined:

$$\operatorname{grad} Q = \sum_{j=1}^{n} \frac{\partial Q}{\partial x_j} \bar{i}_j. \tag{3}$$

Here i_j are the unit vectors in the space $x_1, \ldots, x_n$, where the x_i are assumed to be rectangular coordinates. After the gradient is found at the given point, a step is made in the direction opposite to the direction of the gradient, i.e., in the direction of the quickest decrease in the magnitude of Q. Following this the gradient is again determined at the new point, another step is taken, etc.

b. The method of steepest descent [3] in which, after the gradient is determined, motion occurs in the direction opposite to that of the gradient until the minimum of Q is found along this direction. Then, an analogous cycle is carried out, consisting of the determination of the gradient and the subsequent operational motion, etc.

After considering the possibilities of the different methods with respect to the rate of convergence, as well as the convenience of introducing various restrictions H_j, for the case of appreciable distances from the minimum, preference was given to the method of steepest descent. This method has the following advantages: a) speed of convergence, especially significant when the number of variables x_i is large; b) realizable by means of a relatively simple circuit; c) convenience of imposing any type of restriction.

The last condition is very important. For example, in the application of the method of successive changing of variables (the Gauss-Zeidel method) the circuit is simple only if there are no restrictions. Their presence greatly complicates the circuit.

The imposition of restrictions is convenient if the gradient method is used (see [4]), but this method takes second place to the steepest descent method on the basis of the speed of convergence. Because of this it is better to use the steepest descent method where large deviations from the minimum are involved. However, for small deviations, the gradient method yields the minimum, stable, mean deviation from the minimum and guarantees better tracking of the slowly shifting minimum. Because of this, when small deviations from the minimum are involved, operations in the apparatus being described are based on the gradient method. Circuits for the method of steepest descent and the gradient method are similar, and transition from one method to the other is no problem. We can find analogies to such a combination of methods in some automatic control systems where different laws of motion are selected for large and small deviations so as to obtain a combination of short regulation time in the transient process, and small error in the steady state.

3. Versions of the Optimalizer without Restrictions

As an illustration of the machine's operation, when large distances from the minimum are involved, let us examine the simplest case of Q as a function of two variable x_1 and x_2 (see Fig. 2 where the (x_1, x_2) plane is shown). Let the initial values of x_1 and x_2 correspond to the point M_0 on this plane. Each cycle consists of a

* Incidentally, this condition is not necessary. It is sufficient that discontinuities of the first kind do not exceed some limiting value.

'trial' motion followed by a 'working' motion. The trial motion consists of a series of consecutive steps Δx_i in the direction of each of the variables separately, and the determination of the components $(\partial Q/\partial x_i)$ of the gradient grad Q at the point M_0. The working motion is in the direction opposite to that of the gradient. At the same time the variables x_i change by discrete increments

$$\Delta x_i = -a\left(\frac{\partial Q}{\partial x_i}\right)_0, \tag{4}$$

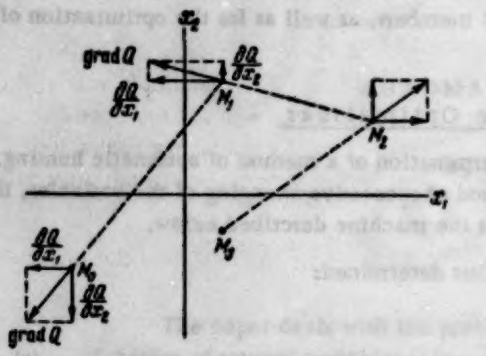


Fig. 2

where a = const > 0. In this way, a transition occurs from point M_0, along the straight line M_0M_1, to point M_1 where the function takes on its minimum value along the given straight line [5]. At this point another trial movement occurs, the gradient is determined, and this is followed by a new working motion along the straight line M_1M_2, etc. This continues until the system arrives close to the minimum and, for example, the quantity $\xi = \sum_{i=1}^{n} |\partial Q/\partial x_i|$ turns out to be smaller than some preassigned, small constant quantity ξ_1. If $\xi_1 < \xi < \xi_2$ (where ξ_2 is a fixed quantity) then the automatic search can be carried out in small steps with small values of a (accurate search). If $\xi > \xi_2$, then the search is carried out in large steps (coarse search). Various sizes of steps make it possible to obtain various combinations of accuracy and speed of the automatic search. Each step, whether trial or work, is accompanied by a measurement, and occurs over a time interval called the access time.

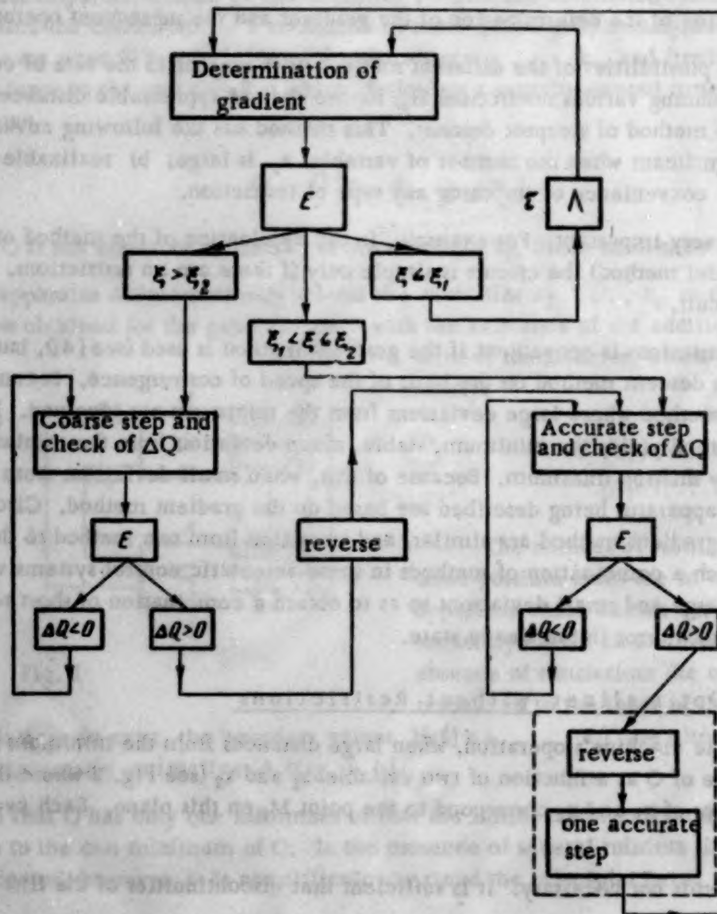


Fig. 3

There are two possible versions of the system:

a. alternative coarse and precise scans are made along each line M_iM_{i+1}: having located the minimum approximately, the system uses small steps to locate it more precisely;

b. a constant size of step is used, determined from the value of ξ before the commencement of each work motion.

The principle of operation of version "a" is clarified with the help of a block diagram of the sequence of operations, shown in Fig. 3. The operation of determining the gradient, with which the machine's work cycle starts, is shown at the top of the block diagram. This is followed by a selection of one of the next possible directions of the operation (operation E). If, for example, $\xi > \xi_2$, then the next operation is a coarse step; if $\xi_1 < \xi \leq \xi_2$, then an accurate step is made; finally, if $\xi \leq \xi_1$, then no working motion ensues and, after some time τ, the operation of determining the gradient is repeated.

Each coarse step is followed in turn by a selection – operation E, depending on the sign of the increment $\Delta Q^{(i)}$ – of the magnitude of the i-th step in question, where

$$\Delta Q_{(i)} = Q_{(i)} - Q_{(i-1)}. \tag{5}$$

Here $Q_{(i)}$ is the magnitude of Q after the i-th cycle, while $Q_{(i-1)}$ is the same quantity after the (i – 1)-th cycle. If $\Delta Q < 0$ then the coarse step is repeated; if $\Delta Q > 0$ then a reversal occurs and subsequent steps are accurate ones in the direction opposite to the initial one. Operation E is repeated after each accurate step. If $\Delta Q < 0$ then the accurate step is repeated; if $\Delta Q > 0$ then a reversal and one accurate step in the new direction (which now coincides with the initial direction of the coarse hunt) are made. With this the cycle for locating the minimum ends. After a time τ a determination of the gradient is repeated.

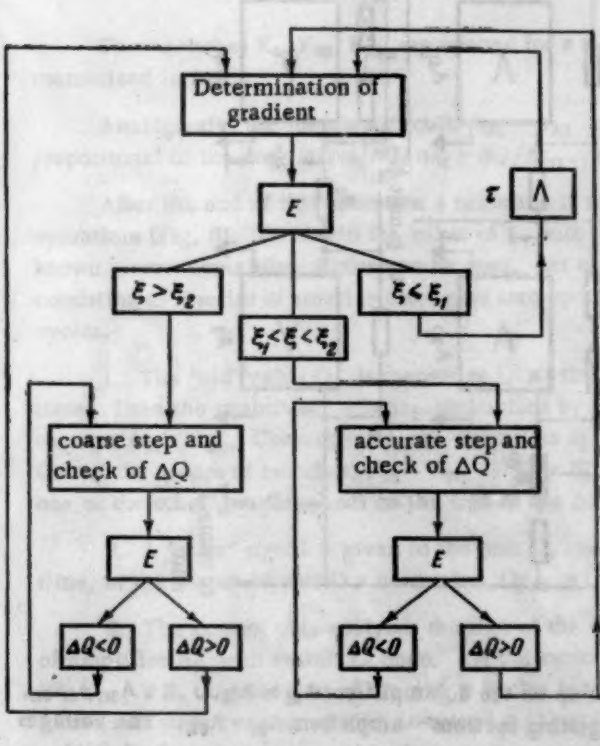


Fig. 4

The circuit may be simplified by omitting the last two steps in the cycle enclosed by a broken line; however, the accuracy of the search is reduced in this case.

Figure 4 is a block diagram showing the sequence of operations in version "b". It is simpler than the preceding one since here there is no reversal. The size of step on each line M_iM_{i+1} depends on the value of ξ which emerges from the operation of determining the gradient. Otherwise the designations and nature of the operations are here the same as in Fig. 3, and therefore do not require special explanation. The search accuracy in this system, for a given number of cycles, is smaller than in the preceding one.

The automatic optimalizer consists of two parts – the operational and control units. The first can be realized using analog-type calculating devices (abbreviated CD), the second using digital-type CD. Thus, the automatic optimalizer appears as a combination of a digital and analog computer. However, no digital-to-analog, or vice versa, converters are necessary since the analog section represents a closed loop which is the operational unit. The digital section – control unit – receives at its input binary quantities (i.e., quantities having only two possible values) from the analog section and, in turn, controls the analog section by means of switches. The principle of operation remains unchanged, irrespective of what type of equipment is used – electromechanical, electronic, etc. As an example, Fig. 5 shows the schematic of a completely electronic operational unit for the "a" version of equipment.

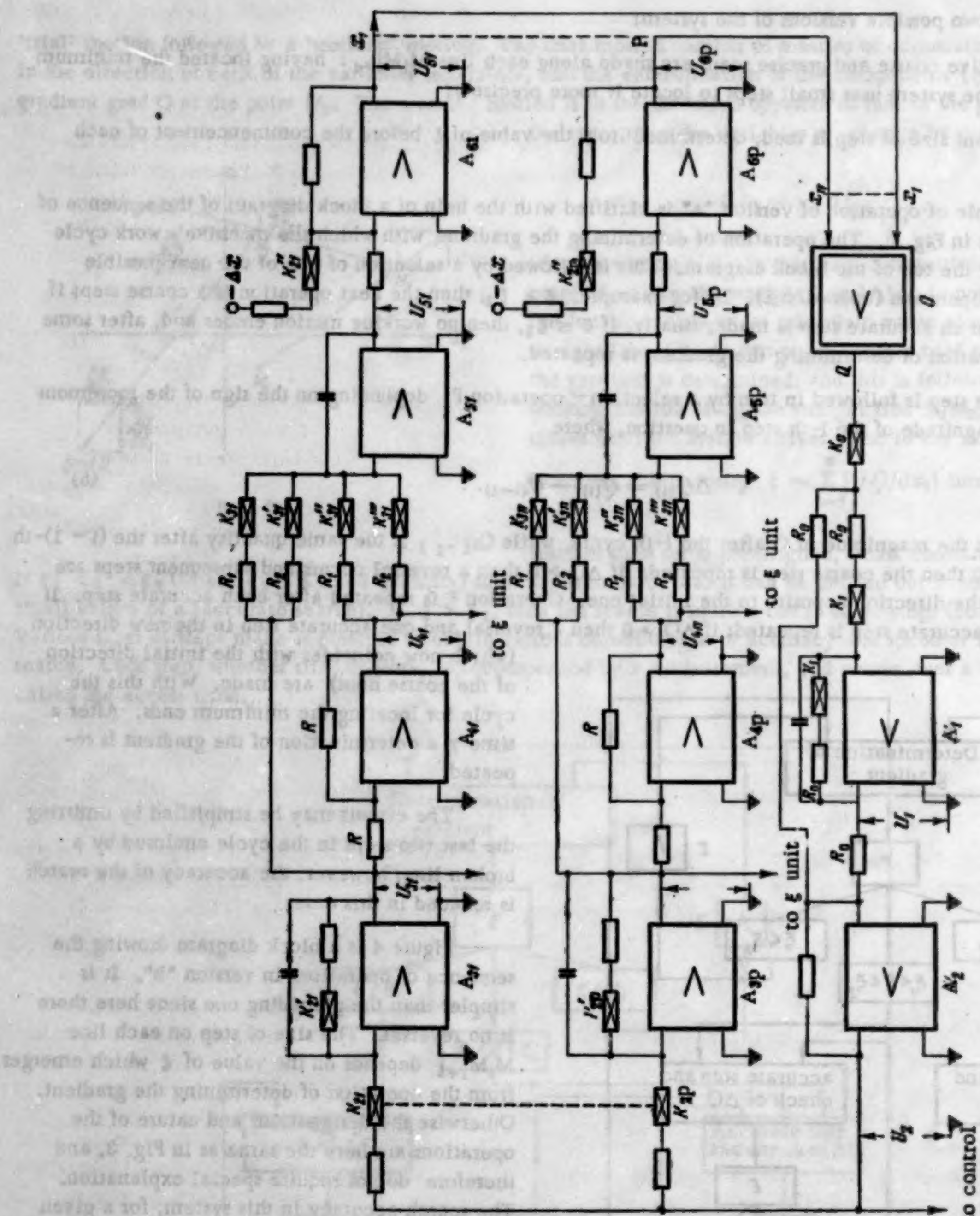


Fig. 5.

Before the unit is put into operation, the zeros are set up on the dc amplifiers A_{31} - A_{61}, . . . , A_{3n} - A_{6n}, A_1 and A_2, and the initial conditions are set up on the integrating sections — amplifiers A_{51} - A_{5n}. The voltages U_{51} - U_{5n} at the output of these sections represent, on a particular scale and with a reverse sign, the initial values of x_i, which are delivered to the output of unit 0 through the inverters A_{61} - A_{6n}. After the zeros and initial conditions are set up, the start button (on the control unit) is pushed and the circuit begins to operate cycle after cycle. A large cycle begins with the operation of determining the gradient. This operation in itself consists of a series of access cycles.

1. Memorizing the "old" value of Q. For this, the electronic switch K_0 at the output of the unit O, and switches K_1 in the circuit of amplifier A_1, which are normally closed, open for a time τ and then reclose.*

*** The opening and closing of the switches is done by the control unit.**

At the output of the unit O there is already, at this time, a value of Q corresponding to the values of x_i delivered to the input. If switches K_1 are open, which can be represented by short circuits,* then section A_1 is inertial. Consequently, its output voltage will become, after a known time τ_0, equal to the quantity Q with a negative sign, i.e., equal to $Q_0 = -Q(x_1, \ldots, x_n)$.

2. Memorizing the partial derivative $\partial Q/\partial x_1$. Here switch K_{21}'', through which a small increment Δx is transmitted to the input x_1 of unit O, is closed. Unit O receives a 'start' signal from the control unit. The unit goes through a work cycle (for example, a transient process occurs in the electronic model) as a result of which there appears at the output of this unit, after some time, the value $Q_1 = Q(x_1 + \Delta x, x_2, \ldots, x_n)$. Now switches K_1 are closed but switches K_0, K_{21}, $K_{21}'^{1}$ are opened for a time τ. As a result, A_2 develops at its output the voltage

$$U_2 = Q_0 - Q_1 = Q(x_1, x_2, \ldots, x_n) - Q(x_1 + \Delta x_1, x_2, \ldots, x_n) = -\Delta Q_1. \qquad (6)$$

When switches K_{21} and K_{21}' are open, section A_{31} represents an inertial section with an amplification factor equal to unity. Therefore, the voltage $U_{31} = \Delta Q_1$ is established at its output and is subsequently memorized there when K_{21} and K_{21}' are closed. In view of the constant nature and small value of the increment Δx, the voltage U_{31} is approximately proportional to $\partial Q/\partial x_1$.

3. Memorization of the partial derivative $\partial Q/\partial x_2$. This is done in an analogous manner. Here switch K_{22}'' is opened, the increment Δx is delivered to input x_2, as a result of which A_2 develops at its output the voltage

$$U_2 = Q_0 - Q_2 = Q(x_1, x_2, \ldots, x_n) - Q(x_1, x_2 + \Delta x, x_3, \ldots, x_n) = -\Delta Q_2. \qquad (7)$$

Then switches K_0, K_{22}, K_{22}' are opened for a time τ and the quantity ΔQ_2, proportional to $\partial Q/\partial x_2$, is memorized in A_{32}.

Analogously, the memory circuits A_{31} - A_{3n} are caused to memorize the values ΔQ_1 - ΔQ_n, which are proportional to the derivatives $\partial Q/\partial x_1$ - $\partial Q/\partial x_n$. With this the operation of determining the gradient ends.

After the end of this operation a selection is made, from the value of ξ, of the nature of the subsequent operations (Fig. 3). To obtain the value of ξ, with various weighting coefficients,** for $|\partial Q/\partial x_i|$, a well known circuit containing diodes can be used. Let us assume, for example, that $\xi > \xi_2$. Then a coarse cycle, consisting of a series of small cycles, goes into operation. Each small cycle consists of the following access cycles.

1. The 'old' value Q_i is memorized. At the same time all the switches K_{31} - K_{3n} open for a certain time. Then the quantities $\partial Q/\partial x_i$ multiplied by certain coefficients, proceed to the inputs of the integrating sections A_{61} - A_{6n}. Consequently, the quantities x_i receive incremental changes determined by Formula (4). One of the groups of switches, K_{31} - K_{3n} or K_{31}'' - K_{3n}'', can be eliminated from the circuit. The operation of one or the other group depends on the sign of the Δx delivered to the input of O.

2. A "start" signal is given to the unit O, and a work cycle occurs. As a result there appears, after some time, at the output of unit O a new value $Q_{i+1} = Q(x_{1i} + \Delta x_{1i}, x_{2i} + \Delta x_{2i}, \ldots, x_{ni} + \Delta x_{ni})$.

3. The control unit analyzes the sign of the difference $\Delta Q_{i+1} = Q_{i+1} - Q_i$, which is formed at the output of amplifier A_2, with switch K_0 open. All the switches in the memory circuit A_{31} - A_{3n} are closed at this time. If $\Delta Q_{i+1} < 0$, then the first of the above cycles is repeated. If $\Delta Q_{i+1} > 0$, then (Fig. 3) a reversal occurs and subsequent steps become smaller in size, i.e., they become "accurate".

In the circuit this is achieved very simply. Instead of switches K_{31} - K_{3n} opening simultaneously with every step, switches K_{31}'' - K_{3n}'' now open simultaneously and deliver to the integrating sections voltages opposite in sign to the former ones. If the new input voltages are delivered through larger resistances ($R_2 > R_1$) then the changing currents of the capacitors in the integrating loops will be smaller. Hence, the voltage increments at the outputs of the integrating loops A_{61} - A_{6n}, also decrease by a corresponding number of times, i.e., steps Δx_k

* In the laboratory model, series-type, diode, two-pole switches, controlled by triggers, were used.

** If weighting coefficients need be taken into account.

will become smaller. Resistors R_1 and R_2 can be made equal but, to obtain a large stop, two switches K_{31} - K'_{31}, for example, would be opened simultaneously, while for a small step only one of these keys would be opened.

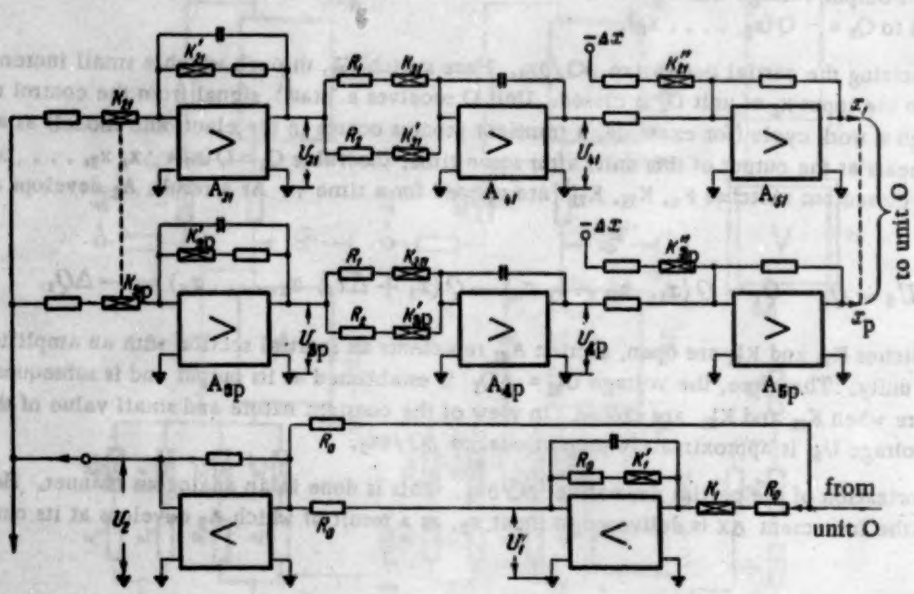


Fig. 6

The same cycles are repeated for an accurate search that are used in a coarse search. When ΔQ_{i+1} becomes greater than zero, then a reverse and one accurate step backwards must follow. This is achieved, for example, by switching on, for one cycle, switches K''_{31} - K''_{3n}, which deliver, through larger resistors R_3, to the inputs of the integrating sections, the voltages U_{41} - U_{4n} which are opposite in sign to the corresponding voltages U_{31} - U_{3n}, respectively.

At the end of the operations described above the gradient is determined again and a new large cycle begins.

Let $\xi \leq \xi_1$. In this case the control unit does not begin a new large cycle, but, after some time, repeats the operation of determining the gradient.

The described principle of operation makes it possible to obtain not only two grades of search steps, but also any number of such grades, for example 'coarse', 'medium', and 'accurate' steps. The principle of achieving such a circuit is apparent from what has been said. It is also possible to control the step gradually depending on the value of ξ. However, experiments have shown that there is no real need for this.

Figure 6 shows a simpler schematic of the basic unit, corresponding to the "b" version and the block diagram in Fig. 4. Since there are no reversals here, invertors A_{41} - A_{4n}, which were present in the first version (Fig. 5), can be omitted.

As in the block diagram in Fig. 4, the large cycle here begins with the determination of the gradient. As a result of this operation, voltages U_{31} - U_{3n} become proportional to the partial derivative $\partial Q/\partial x_1$ - $\partial Q/\partial x_n$, respectively.

The analysis of the quantity ξ is carried out in the same way as in the preceding circuit. If this is followed by an accurate search, then switches K_{31} - K_{3n} are simultaneously switched on for a certain length of time, if it is a coarse search then switches K'_{31} - K'_{3n} are also switched on. Let us note that in one channel it is sufficient to have only one switch, say K_{31}, but it must be turned on for a longer time if a coarse search is required, and a shorter time if an accurate search is required. In this case the circuit of the operational unit is somewhat simplified, but the control unit circuit becomes more complicated.

The analysis of the sign of ΔQ_i can be carried out, for example, by means of a simple circuit, the output of which can have only two values: upper (nominally 1) and lower (nominally 0). When $U_2 = -\Delta Q_i > 0$, the output of the circuit is in state 0. When $U_2 = -\Delta Q < 0$, the output changes to state 1.

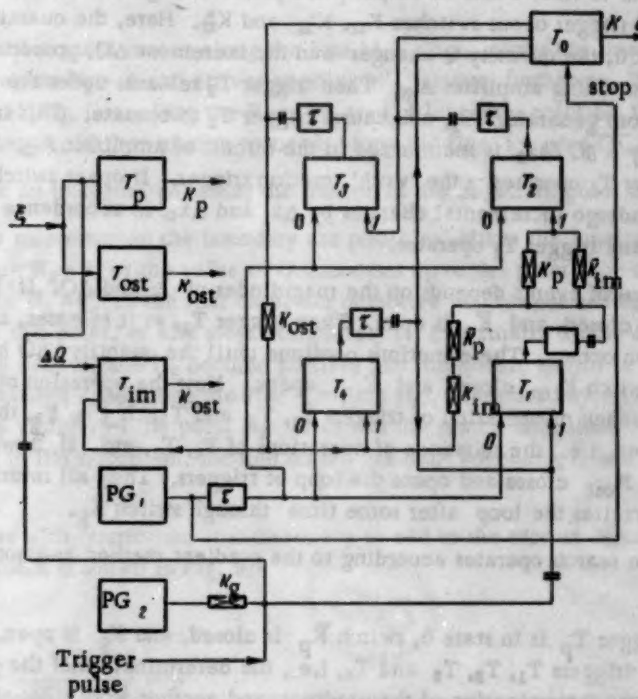


Fig. 7

Switches K'_{31} - K'_{3n} can be omitted from the circuit in Fig. 6, and one grade of search step is sufficient because the automatic optimalizer includes an automatic step control. In fact, it can be seen from Equation (4) that during the work motion, the increment Δx_i is proportional to $|\partial Q/\partial x_i|$. For appreciable distances from the minimum, i.e., where $|\partial Q/\partial x_i|$ is usually large, the size of the step Δx_i is also large. Close to the minimum, where $|\partial Q/\partial x_i|$ is small, the size of the step also decreases. Because of this the laboratory model of the apparatus did not contain a changeover from a "coarse" to an "accurate" search although, in some cases, such a transition is extremely desirable. The quantity ξ was used for another purpose — for the changeover, for small ξ, from minimization by the method of steepest descent to minimization by the gradient method.

Figure 7 shows the block schematic of one version of the control unit. For the sake of simplicity, the control unit for the simplified circuit in Fig. 6, for $n = 2$, is examined. However, analogous principles can be applied to the design of a control circuit for a more complex system, shown in Fig. 5.

Triggers T_1, T_2, T_3 and T_4, are connected in a closed loop. When trigger T_1 is transferred from state 1 to state 0, its output pulse passes through the delay circuit τ and through either switch K_p or $\overline{K}_{im}$, and transfers trigger T_2 from state 0 to state 1. Trigger T_2, in its turn, returning from state 1 to state 0, switches over trigger T_3 from 0 to 1. In the same way, trigger T_3 acts, through switch K_{ost}, on T_4, and T_4 acts on T_1. Besides the large loop T_1, T_2, T_3 and T_4, a small closed loop can be formed from T_1 and T_4, if switches $\overline{K}_p$ and K_{im} are open.

Trigger T_0 gives the 'start' signal to unit O, when it changes from 0 to 1, and the 'stop' signal, when it returns from 1 to 0. Trigger T_p operates (changes to 1) for large ξ ($\xi \leq \xi_2$), while trigger T_{ost} operates for small values of ξ ($\xi \leq \xi_1$). When $\xi_1 < \xi < \xi_2$ triggers T_p and T_{ost} are in the 0 state.

Trigger T_{im} operates when $\Delta Q > 0$ and is in state 0 when $\Delta Q < 0$.

Let us trace the sequence of operations in the control unit. Switch K_g is in the output of the pulse generator PG_2 which supplies low frequency pulses. Switch K_g is only open when all the triggers T_1 - T_4 are in the 0 state.

In this case (as well as when a trigger pulse appears) the pulse arrives at the 1 input of trigger T_1 and transfers it into state 1. As trigger T_1 operates, it opens switches K_1, and K'_1 (Fig. 6) as a result of which a value of Q is memorized. After some time, a pulse from generator PG_1, operating at an appreciably higher frequency than PG_2 (these generators do not have to be synchronized with each other), appears at the 0 input of trigger T_1. Consequently, T_1 switches over to state 0 and its output pulse passes through either of the switches K_p and $\overline{K}_{im}$, and operates trigger T_2. This trigger opens switches K_{21}, K'_{21} and K''_{21}. Here, the quantity x_1, with increment Δx is fed to the input of unit 0, the quantity Q changes and the increment ΔQ, proportional to $\partial Q/\partial x_1$, is memorized by the section containing amplifier A_{21}. Then trigger T_2 releases under the action of the next pulse, which arrives at its 0 input from generator PG_1, and causes trigger T_3 to operate. This trigger opens gates K_{22}, K'_{22} and K''_{22}, and the quantity $\partial Q/\partial x_2$ is memorized in the circuit of amplifier A_{22}. Then trigger T_{im} is put into the 0 state and trigger T_4 operates — the 'work' motion trigger. It opens switches K_{31} - K_{32} (Fig. 6), because of which x_1 and x_2 undergo incremental changes by Δx_1 and Δx_2, in accordance with Equation (4). After one step, trigger T_4 releases and trigger T_1 operates.

The subsequent sequence of events depends on the magnitudes of ξ and ΔQ. If $\xi \geq \xi_2$, then trigger T_p is in the 1 state, switch K_p is closed, and $\overline{K}_p$ is open. Then trigger T_1, as it releases, causes trigger T_4 to operate. Another work motion occurs. These motions continue until the quantity ΔQ becomes greater than zero. Then T_{im} operates, switch K_{im} closes, and $\overline{K}_{im}$ opens. Now the operation of determining the gradient begins, i.e., the sequence of operation of triggers T_1, T_2 and T_3. If $\xi \geq \xi_2$, then this operation is followed by a new work motion, i.e., the sequence of operations of T_4, T_1, etc. If, however, $\xi \leq \xi_1$, then trigger T_{ost} operates, switch K_{ost} closes and opens the loop of triggers. They all return to state 0, but subsequently, a pulse from PG_2 excites the loop after some time through switch K_g.

If $\xi_1 < \xi < \xi_2$ then the search operates according to the gradient method and not the method of steepest descent.

In fact, in this case trigger T_p is in state 0, switch $\overline{K}_p$ is closed, and K_p is open. This is followed by the consecutive excitation of triggers T_1, T_2, T_3 and T_4, i.e., the determination of the gradient, and one work step. This is followed by a new determination of the gradient and another work step, etc. Trigger T_0 operates in the case when at least one of the triggers T_2, T_3 and T_4 switches over to the 1 state. Then the 'start' signal is given to the unit 0 which can be, for example, in a particular case, an electronic unit. The next pulse from generator PG_1 returns trigger T_0 to the 0 state ("stop" signal) before triggers T_2, T_3 and T_4 switch over to the 0 state. Triggers T_1, T_2, T_3, T_4 and T_0 can be replaced by monostable flip-flop oscillators. This simplifies the circuit, since generator PG_1 is no longer required — the circuits return themselves to the 0 state. However, this is accompanied by complications in the resetting of the unit's operating speed.

4. Automatic Optimalizer with Restrictions

The principle of operation of this optimalizer is illustrated in Fig. 8 for the case of two variables x_1 and x_2 and two restrictions $H_1 \leq 0$ and $H_2 \leq 0$. However, this principle is applicable with no change to the case of any number $\underline{n}$ variables and an arbitrary number $\underline{m}$ restrictions.

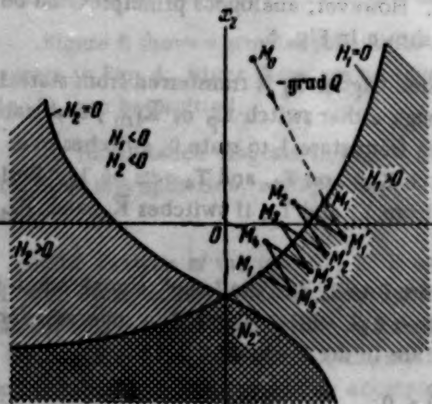


Fig. 8

The lines $H_1 = 0$ and $H_2 = 0$ are plotted on the (x_1, x_2) plane, and the 'forbidden' regions $H_1 > 0$ and $H_2 > 0$ are shaded. While the objective point (x_1, x_2) moves in the permitted, unshaded region, the system operates in the same way as the one discussed before. However, if even one of the quantities H_j becomes greater than zero, the system's operation changes. The machine's 'attention' switches from the quantity Q to those of the quantities H_j which have become greater than zero, and their sum undergoes the same operations which previously the quantity Q underwent in the permitted zone. Suppose, as an example, the quantities $H_{j1}, H_{j2}, \ldots, H_{jl}$ $(l \leq m)$ become greater than zero. Let us denote

$$H_0 = H_{j1} + H_{j2} + \ldots + H_{jl}.$$

Now the operation of determining the gradient proceeds as follows: using the same method as before, $\partial H_0/\partial x_i$ ($x_i = 1, \ldots, n$) are calculated and are memorized at the outputs of the same sections $A_{31} - A_{3n}$ (Figs. 5 and 6) as before. Then the objective point moves along a line (− grad H_0) until all the quantities H_{jk} ($k = 1, \ldots, l$) become negative. As soon as this is achieved, the unit returns to analyzing the quantity Q alone.

In Fig. 8 the objective point moves along the line M_0M_1, i.e., along the line (− grad Q). In the process it enters the forbidden region $H_1 > 0$ and arrives at point M'_1. Hence, further on, H_1 is determined and the point proceeds along the line M'_1M_2, i.e., along the line (− grad H_1). At the point M_2 the magnitude of H_1 becomes less than zero. Consequently, further motion proceeds along the line (− grad Q), etc.

Figure 8 shows, on an exaggerated scale, the motion of the objective point along the boundary $H = 0$.

Various cases of a minimum on the boundary are possible: either the function Q has a minimum at the point N_1, on the boundary $H_1 = 0$, or the value of Q decreases up to the point N_2 — the points of intersection of the boundaries $H_1 = 0$ and $H_2 = 0$. In the first case, the objective point, upon reaching N_1, begins to perform small oscillations about this point, or else stops completely if ε is small. In the second case the point crosses the line $H_2 = 0$, in which case H_1 and H_2 become positive and the system begins to act on the quantity $H_0 = H_1 + H_2$, and leaves the forbidden zone along the line (− grad H_0). Here the objective point begins to perform small oscillations in the vicinity of the point N_2. With this the search ends since N_2 is the required point. At the point corresponding to the minimum, located on the gradient boundary, Q and H_0 are directed toward each other.

For an optimalizer with restrictions it is necessary to add to the circuit described earlier an additional device, the circuit of which is shown in Fig. 9.

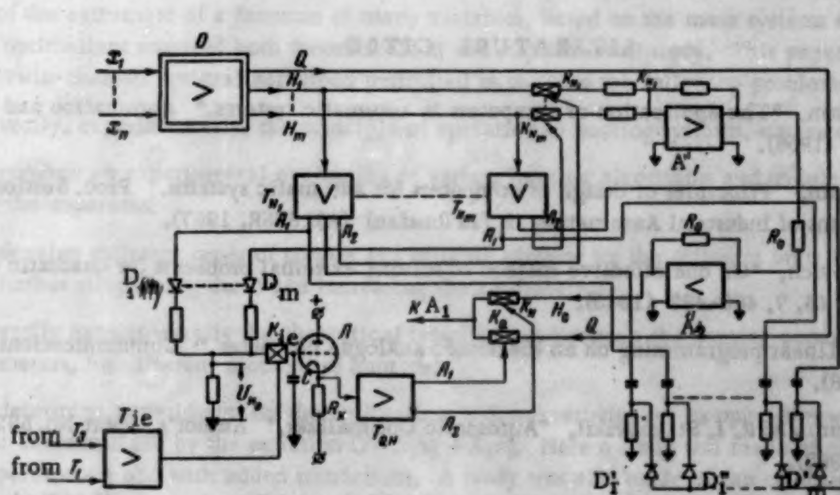


Fig. 9

When any of the quantities H_j, at the output of the unit 0, become positive, the triggers corresponding to them, from the set $T_{H_1} - T_{H_m}$, operate. Dc amplifiers can be introduced in front of these triggers to increase the system's sensitivity to the sign of H_j. When trigger T_{H_j} operates, it opens switch K_{H_j} at the input of amplifier A^*_1. As a result, invertor A^*_1 develops at its output the necessary sum H_0, in accordance with Equation (8). The operated triggers also act on the diode circuit $D_1 - D_m$ (an "Or" - type logic circuit) at the output of which appears the positive voltage U_{H_0}. The voltage U_{H_0} acts through the switch K_{b1} and the cathode-follower on the trigger T_{QH}, which opens the switch K_H and closes switch K_0. Now H_0, and not Q, is passed to the input of amplifier A_1 (Fig. 5). Beside this, a negative pulse, from any anode of any of the triggers T_{H_j}, proceeds through the diodes D'_1, $D''_1 - D'_m$, D''_m to the input of the trigger T_{ost} (Fig. 7) and puts the control unit into the "determination of gradient" condition. Let us assume that the system begins to operate with the

quantity H_0. As soon as the operation of determining the gradient begins and trigger T_1 operates (Fig. 7), it causes T_{bl} to operate (Fig. 9). The latter closes switch K_{bl}, and the state of trigger T_{QH} "freezes" until the end of the operation of determining the gradient. At the end of this operation trigger T_3, in releasing, switches trigger T_{bl} to the 0 state, and switch K_{bl} opens. As soon as all the quantities H_j become negative, all the triggers T_{H_1} - T_{H_m} release (return to the 0 state), the quantity U_{H_0} becomes zero and, after K_{bl} is opened, releases trigger T_{QH}. Now the quantity Q is once again passed to the input of amplifier A_1. The pulse from the triggers T_{H_j} causes trigger T_{ost} to operate. In this way, the circuit is once again returned to the condition of "determination of gradient."

There are other methods of including the restrictions, with which the search rate along the boundary is increased somewhat. However, this is accompanied by appreciable complications in the circuit of the restrictions unit (Fig. 9).

For large values of n it is convenient to introduce certain changes in the optimalizer, permitting thereby a great economy in components. The problems of designing such systems, for large values of n, must appear as the subject of a separate paper.

The systems considered above are not universal. In practice there arises the possible case where, instead of an isolated maximum in the space $(x_1, x_2, \ldots, x_n)$ there is an $(n-k)$-dimensional space, where $k = n-1, \ldots, 1$, in which Q has the same maximum value [6]. In this case the logic of the system's operations to achieve such a sub-space must be more complicated; here a much larger number of trial movements becomes necessary. However, experiments have shown that even the relatively simple logic described above, leads to success in many important examples. The problems of complicating the logic for the special cases mentioned above, as well as to allow for the effect of variations in the function $Q(x_1, \ldots, x_n)$ itself during the time of the trial motion, if this does occur, are the subject of a separate paper.

LITERATURE CITED

[1] A. A. Fel'dbaum, "The application of computers in automatic systems," Automation and Remote Control (USSR) 17, 11 (1956).

[2] A. A. Fel'dbaum, "Principles of design of computers for automatic systems," Proc. Session AN SSSR on the Scientific Problems of Industrial Automation 2 [In Russian] (AN SSSR, 1957).

[3] L. V. Kantorovich, "On one effective method of solving extremal problems for quadratic functions," Dokl. AN SSSR (USSR) 48, 7, 483-487 (1945).

[4] J. B. Pyne, "Linear programming on an electronic analogue computer," Communications and Electronics May (1956).

[5] A. A. Fel'dbaum and R. I. Stakhovskii, "Automatic Optimalizer," Author's Report No. 582931/26 [In Russian] (Sept. 5, 1957).

[6] O. B. Davies, in the book: The Design and Analysis of Industrial Experiments (New York, 1954).

Received December 12, 1957

TWIN-CHANNEL AUTOMATIC OPTIMALIZER

R. I. Stakhovskii

(Moscow)

The circuit and some units of a twin channel, electronic, automatic optimalizer for the location of minima (taking restrictions into account) are described in detail. Experimental results, corresponding to its different operating modes, are given.

INTRODUCTION

The object of this work, some of the results of which are presented here, was the design of a machine for the location of the extremum of a function of many variables, based on the main systems suggested in [1]. The design of the optimalizer required both theoretical [2] and experimental study. This paper is devoted to the description of a twin-channel optimalizer which permitted us to solve the following problems.

1. To verify, experimentally, the principle of operation of hunting systems, suggested in [1].

2. To conduct an experimental evaluation of various hunting algorithms and to select the best operating algorithm for the apparatus.

3. To develop different types of simple and reliable circuits for the different units in the machine and to plan ways of further simplifying them and increasing their reliability.

4. To verify experimentally the theoretical relationships between the hunting accuracy and time and the system's parameters, for different methods of hunting.

5. To determine experimentally the optimalizer's characteristics for its operation with a standard single capacity plant characterized by the equation $Q = A_1x_1^2 + A_2x_2^2$. Here a study was made of plants with constant and variable parameters and with added restrictions. A study was also made of the condition under which the plant was acted on by random interferences.

Systems for the automatic location of extremums, with or without added restrictions, can be used as the basis of automatic machines designed for a wide variety of applications, beginning with a machine for the solution of variational problems, and ending, for example, with an automatic-operator in a production plant.

The design of a twin-channel type of optimalizer represents only the first step in the development of multichannel systems. Consequently, the scheme was designed with the possibility of it being developed for the case of a function of $\underline{n}$ variables in mind. This resulted in the choice of a circuit with a small number of amplifiers per channel, which makes the scheme adaptable to an increase in the number of variables. To simplify the circuitry it was necessary to select a simple operating algorithm for the optimalizer, from the types described in [1], and a simple block diagram of the control unit, in which the hunting method can be changed during the actual hunting.

Some of the problems listed above are clarified in [1]. Therefore, there is no need to dwell on them here. Let us note only that the methods of simplifying the over-all system, as well as its separate units, are far from exhausted. Further simplifications must be made during the design of multichannel systems.

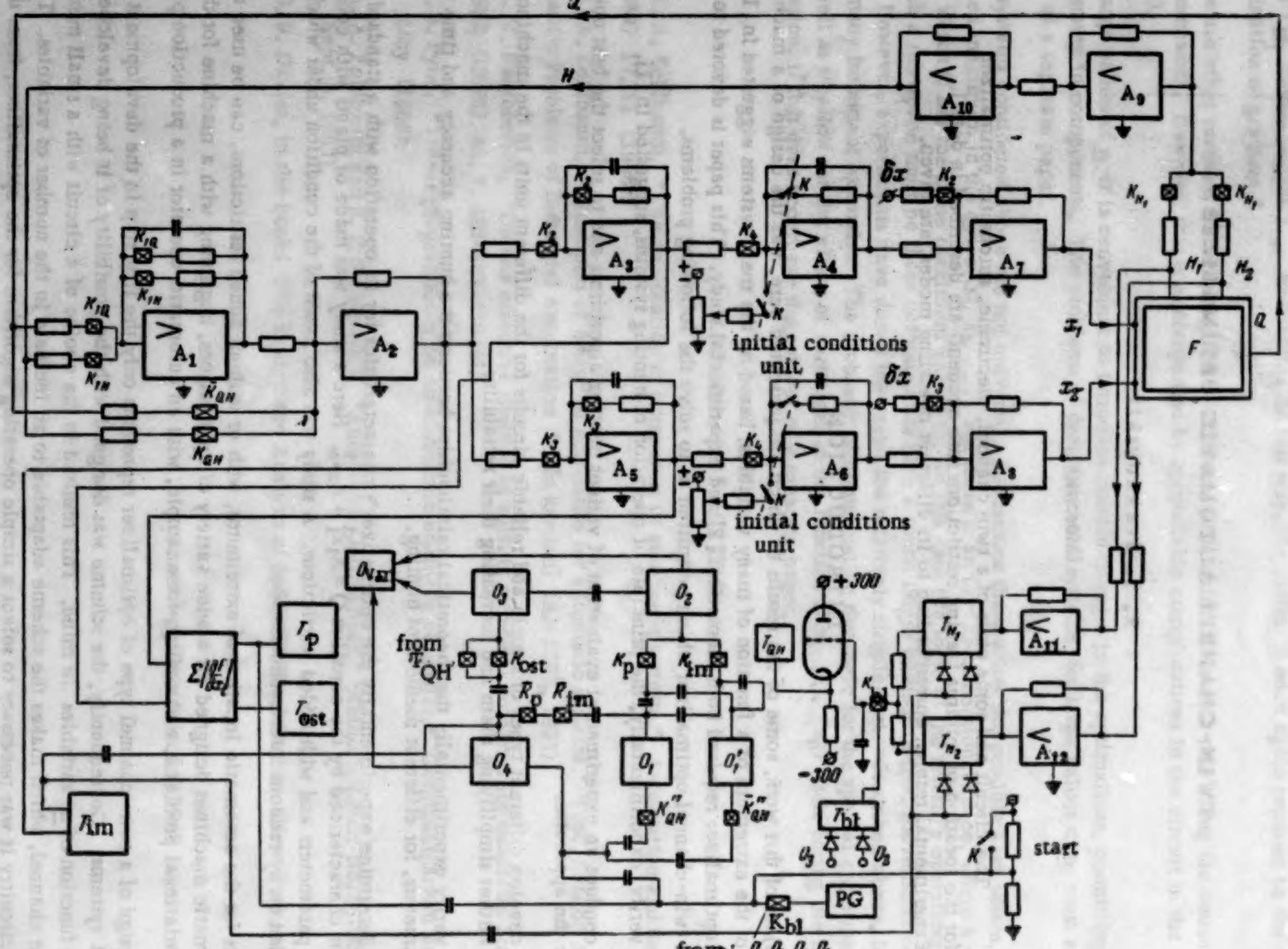

initial conditions unit
initial conditions unit
start
from O_1, O_2, O_3, O_4
from T_{QH}
PG
+300
-300
F
δx
δx
A_1
A_2
A_3
A_4
A_5
A_6
A_7
A_8
A_9
A_{10}
A_{11}
A_{12}
K_{bl}
K_{ost}
K_p
K_{1m}
R_p
R_{1m}
T_p
T_{ost}
T_{bl}
T_{1m}

Fig. 1

General Description of Equipment

The complete schematic of an optimalizer which will hunt when two restrictions are imposed, is shown in Fig. 1. The equipment is divided into three units: operational unit, which is the nucleus of the equipment, control unit and restrictions unit. The details of the block schematic are described in [1] and, therefore, only a general description of the equipment, as well as separate, particular additions and changes, are given below.

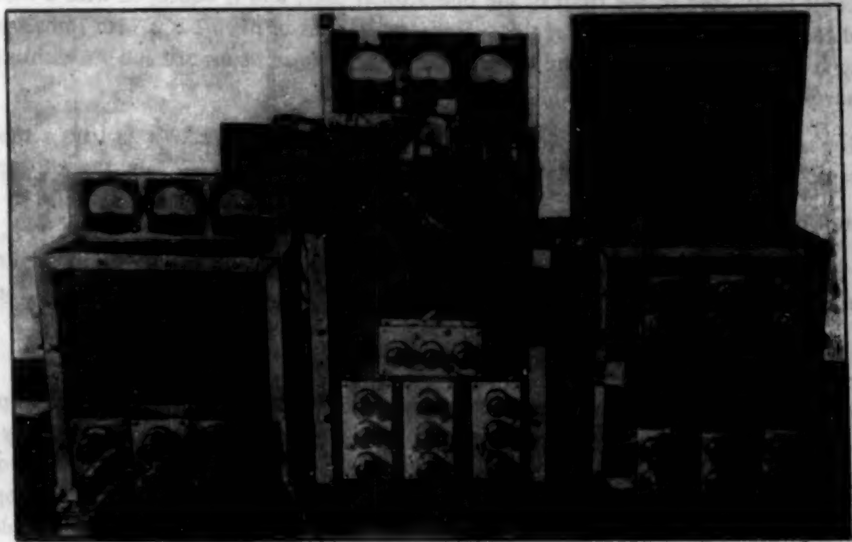


Fig. 2

Since a simplified algorism, in which there is no reversal, is used in this equipment, each channel consists of three amplifiers with the necessary number of electronic switches. The input of the operational unit is a device which determines differences between the values of the input function at two discrete instants of time ("differentiator"). This device has two inputs: for the minimizing and resulting restriction functions H_0. The setting-up of initial conditions is carried out on the step integrators A_4 and A_6. Trial increments are given by means of switches in the summing amplifiers A_7, A_8.

The control unit is designed as a two-circuit loop, composed of the single flip-flop oscillators $O_{1Q}(O_1H)$, O_2, O_3, O_4, which control the group of switches $K_{1Q}(K_{1H})$, K_2, K_3, K_4, in the operational unit.

The switching of the loop circuit, which is necessary when a change is made in the hunting method, as well as when the steepest descent method of hunting is used, is achieved by means of switches K_p, $\overline{K}_p$, K_{im}, $\overline{K}_{im}$, which are controlled by triggers T_p and T_{im}.

To determine the instant at which the hunting algorism is switched, a unit is used which measures the quantity $\xi = \sum_{i=1}^{2} |\partial Q / \partial x_i|$.

The restrictions unit is designed to switch the inputs from Q to H_0, by means of the transfer of control from switches K_{1Q} to analogous switches K_{1H}, if even one of the quantities $H_i(x_1, x_2)$, $i = 1, 2$, would exceed zero. At the same time, the operation of determining the gradient is initiated. In order to avoid false steps, trigger T_{QH}, which switches the inputs, is locked during the determination of the gradient. The unit contains triggers T_{H_1} and T_{H_2} with amplifiers A_{11} and A_{12}. The latter are provided for more precise operation of the triggers. T_{H_1} and T_{H_2} are designed to introduce one or two restrictions by means of switches K_{H_1} and K_{H_2}, as well as to connect in the device for determining the gradient, every time the boundary is crossed. The pulse generator PG serves to ensure the uninterrupted operation of the loop of monostable flip-flops.

Figure 2 shows a photograph of a laboratory model of the optimalizer. The operational unit is in the center. The operational amplifiers are situated in the lower part of the front panel. Above them is the memory circuit

with lugs to which the inputs and outputs of the amplifiers are brought out. The input resistors and those used in the feedback loops are situated in the memory circuit or else in the switching circuits on the rear panel of the unit.

The test panel, with meters x_1, x_2, $Q(x_1, x_2)$, ΔQ, ε, is situated on top. In front of the test panel, and set up on the operational unit, is a model of a nonlinear unit with a quadratic characteristic. Next to it, on the right side, are manual controls for setting up the initial conditions.

The control unit is situated to the right of the operational unit, and is equipped with a memory circuit. This circuit is used to display the operation of the switches by means of neon bulbs with light when the switches are in a conducting state. Various parts of the control unit, in the form of circuit sub-assemblies, are distributed on the front and rear panels of the unit.

The restriction unit, located to the left of the operational unit, is also equipped with a memory circuit for its own operations. Amplifiers A_9, A_{10} and A_{11}, A_{12}, triggers T_{H_1}, T_{H_2}, T_{QH} and the corresponding electronic switches, are situated on the front and rear panels of the unit.

It should be noted that experimentation has shown that the use of display-type memory circuits is very expedient, because it helps in the adjustment and checking of the equipment's operation in different modes.

Short Description of Separate Components

1. Operational amplifier. The requirements for an operational amplifier used in the optimalizer, differ appreciably from the requirements for computing amplifiers used in computers. Actually the requirements for accuracy and permissible amount of zero drift are less stringent, but the reliability requirements are higher, especially if industrial applications are contemplated. Actually, the amplifiers in the optimalizer are servo-amplifiers, and errors in them lead merely to a slowing down of the hunting process, but do not affect the value of the determined minimum. The only exceptions are the amplifiers in the input section of the operational unit, which demand more precise adjustment. However, even here the circuit could be simplified considerably. A very simple and reliable circuit of a three-stage, operational amplifier, using 6N9 and 1/2 6N8 tubes, and deriving its power not from a vibrapack but from regulated power supplies (+ 300, − 190, − 300 v), is utilized in the apparatus. Zero drift over an extended period of time (weeks), does not exceed ± 1.5 v, which turns out to be quite sufficient for the reliable operation of the whole equipment if the range of output voltages is ± 100 v.

2. Electronic switches. Two-pole, series switches are used in the operational amplifiers. The circuit of such a switch is shown in Fig. 3. When a positive voltage, with respect to ground, is delivered to point 1, simultaneously with a negative voltage to point 2, diodes D_1 and D_2 conduct since positive voltages, with respect to cathodes, appear on the anodes of both diodes. The output signal passes through the switch by way of D_1 or D_2, depending on the polarity of the signal. When the controlling voltages on points 1 and 2 are reversed, diodes D_1 and D_2 disconnect the input resistors from the input to the amplifier becuase the anodes of D_1 and D_2 become negative with respect to their cathodes, due to the low forward resistance of D_3 and D_4.

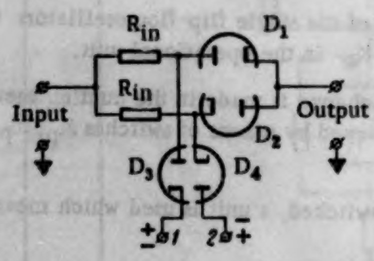
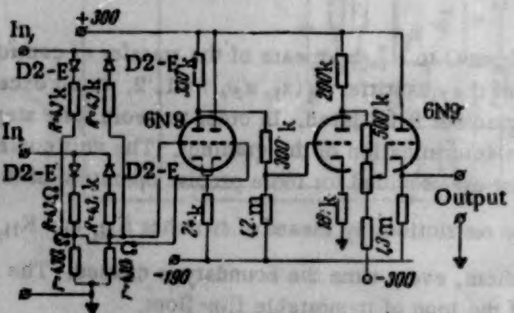


Fig. 3

The switch's most important characteristic, as it is used in the optimalizer, is the small drift of the operational amplifier, with electronic switches in its input, when the amplifier operates as an integrator. The drift is caused by a parasitic emf which exists in the equivalent circuit of the switch, and which depends, all other things being equal, on the cathode temperature and the dissimilarity of the parameters of diodes D_1 and D_2. In a circuit containing an integrating section, which operates from the switch, it is easy to obtain a drift not exceeding (when the section's time constant $R_{in}C = 1$ sec) ± ± (1 - 3) v/min for the open and closed states of the switch.



k ≡ kilohm, m ≡ megohm

Fig. 4

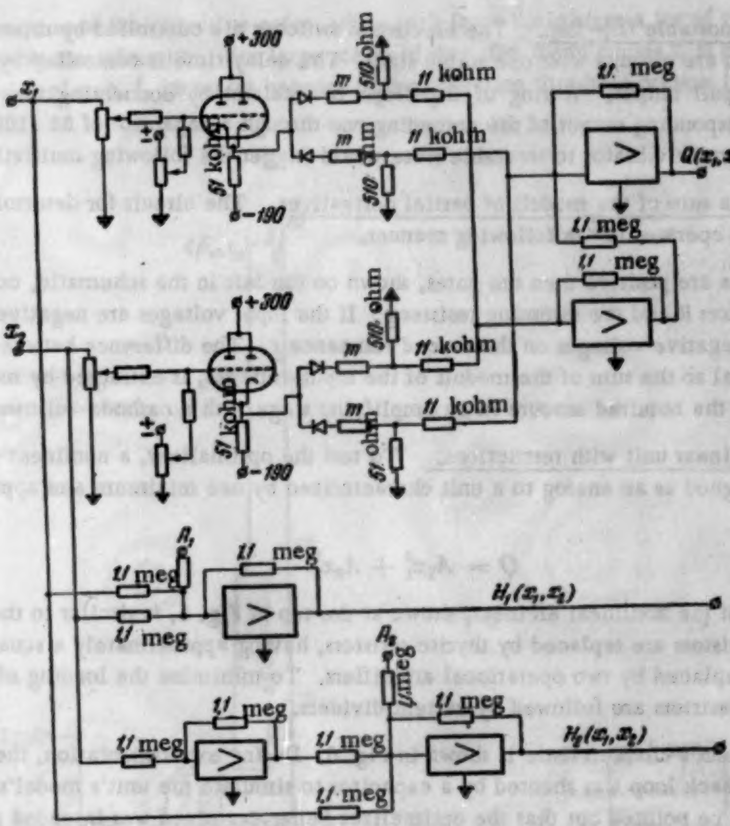

+300
510 ohm
11 kohm
11 meg
51 kohm
-190
$Q(x_1, x_2)$
+300
11 kohm
$H_1(x_1, x_2)$
$H_2(x_1, x_2)$

Fig. 5

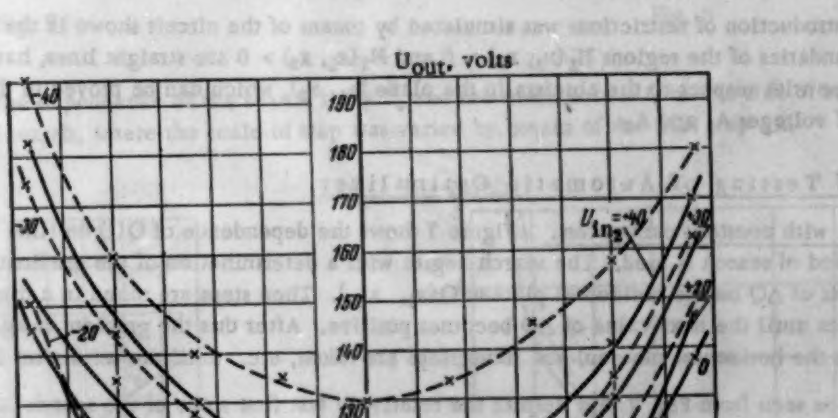

U_{out}, volts
$U_{in_2} = +40$
U_{in_1}, volts
-42 -36 -30 -24 -18 -12 -6 0 6 12 18 24 30 36 42

Fig. 6

3. Controlling monostable flip-flop. The electronic switches are controlled by monostable flip-flops in the control unit, which are circuits with one stable state. The delay time is controlled by the grid bias. The circuit has a push-pull output. A ring of flip-flops is obtained by connecting the input of a multivibrator to the corresponding output of the preceding one through a capacitor of 50 - 100 $\mu\mu$f, whereby the return of one monostable multivibrator to its stable state would trigger the following multivibrator.

4. Finder of ξ as a sum of the moduli of partial derivatives. The circuit for determining ξ is shown in Fig. 4. The arrangement operates in the following manner.

If the input voltages are positive then the gates, shown on the left in the schematic, conduct positive currents through the resistors R and the summing resistor r. If the input voltages are negative then the gates on the right form a sum of negative voltages on the second resistance r. The difference between the voltages on the resistors r, proportional to the sum of the moduli of the input voltages, is extracted by means of a subtracting stage and is amplified by the required amount in an amplifying stage with a cathode-follower output.

5. Model of a nonlinear unit with restrictions. To test the optimalizer, a nonlinear converter, for two input quantities, was designed as an analog to a unit characterized by one minimum and approximated by the equation

$$Q = A_1x_1^2 + A_2x_2^2.$$

Part of the circuit of the nonlinear element, shown at the top of Fig. 5, is similar to the circuit of the ξ_m finder if the input resistors are replaced by thyrite resistors, having approximately a square law characteristic. The subtraction stage is replaced by two operational amplifiers. To minimize the loading of the last stage of the amplifier the thyrite resistors are followed by current dividers.

The nonlinear element's characteristic is shown in Fig. 6. During experimentation, the resistance of the summing amplifier's feedback loop was shunted by a capacitor to simulate the unit's model's time constant $\tau_{out} \approx 0.7$ sec. It should be pointed out that the optimalizer being examined was intended primarily for operation with purely discrete units (units of the 2-nd kind), and the experiments were conducted with a model of a unit of the 1-st kind [3] only for the sake of simplicity.

The introduction of restrictions was simulated by means of the circuit shown in the lower part of Fig. 5. Here the boundaries of the regions $H_1(x_1, x_2) > 0$ and $H_2(x_1, x_2) > 0$ are straight lines, having a positive and a negative slope with respect to the abscissa in the plane (x_1, x_2), which can be moved in this plane by changing the values of voltages A_1 and A_2.

Results of Testing of Automatic Optimalizer

Object with constant parameters. Figure 7 shows the dependence of Q(t) on time when the steepest descent method of search is used. The search begins with a determination of the gradient which is marked by two overshoots of ΔQ on the horizontal plateau $Q(x_{10}, x_{20})$. Then steps are taken in a direction opposite to that of the gradient until the next value of ΔQ becomes positive. After this the gradient is again determined (two overshoots on the horizontal plateau) and more steps are taken, etc. until the minimum is located.

It can be seen from Fig. 7 that despite the relatively fast first stage of the search (already in the first step $Q_{min} \approx Q_{pre}$, insofar as in the object $A_1 \approx A_2$, i.e., the gradient at any point in the plane (x_1, x_2), is in the direction of the minimum) the total search time is nevertheless large, because an appreciable length of time is required to deplete the error which is obtained when the minimum is overshot due to the large size of step.

Figure 8 shows the diagram of a hunt based on the gradient method. Here overshoots occur at each step, and the value Q_{pre} is approached very smoothly. The stabilized value of the mean excess f_0 over Q_{pre} ($f_0 = Q_{mean}$ in the stable state) in the case of the gradient method of search is considerably smaller than for the steepest descent method.

Figure 9 shows the search time T_0, and the mean excess in the stable state f_0 over the magnitude Q_{pre}, as a functions of the scale of step for a steepest descent method of search. The scale of step was varied by beams of the amplification factor of A_2 in the input differentiator. From Fig. 9 it can be seen that the curve T_0 (k) has a minimum. This can be explained by the fact that for a small scale of step the search time is large

because of the slow motion of the objective point, while with large scale steps a lot of time is lost in depleting the errors in determining the minimum. For large scales of step, the mean excess f_0 is practically independent of k. For smaller scales of step, f_0 increases because of the operating threshold present in the trigger for detecting the minimum T_{im}.

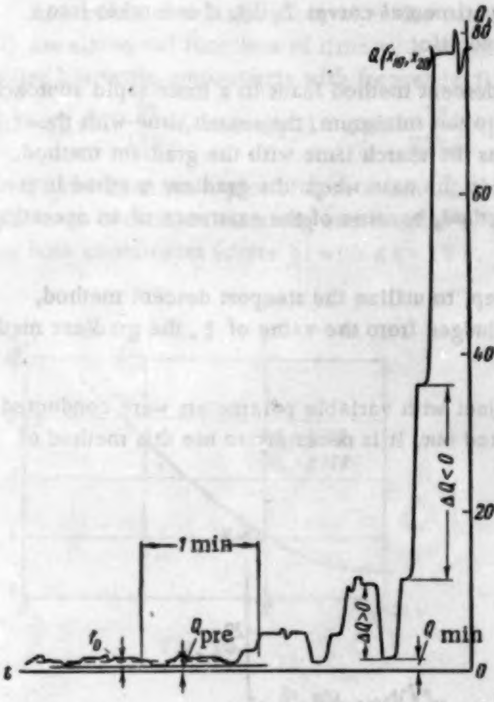


Fig. 7

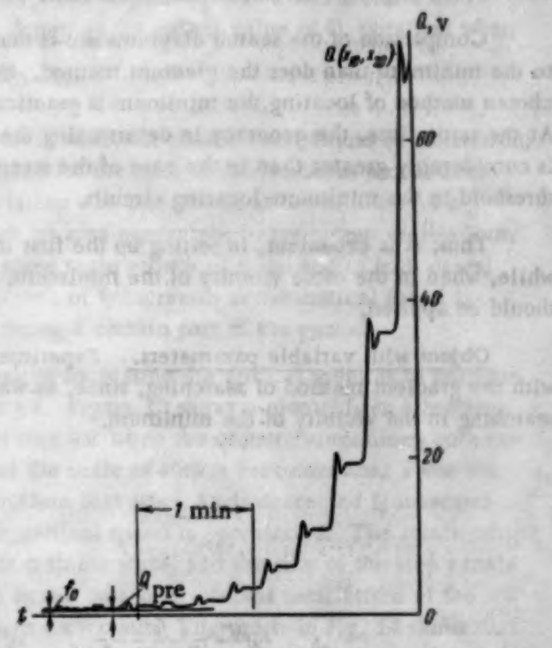


Fig. 8

Figure 10 shows the dependence of the search time T_0 and the mean excess f_0 on the size of step scale for the gradient method of search, where the scale of step was varied by means of the trial step δx.

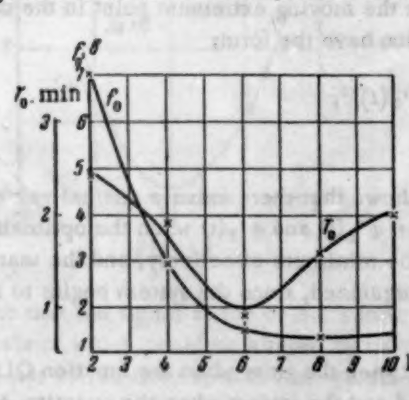


Fig. 9

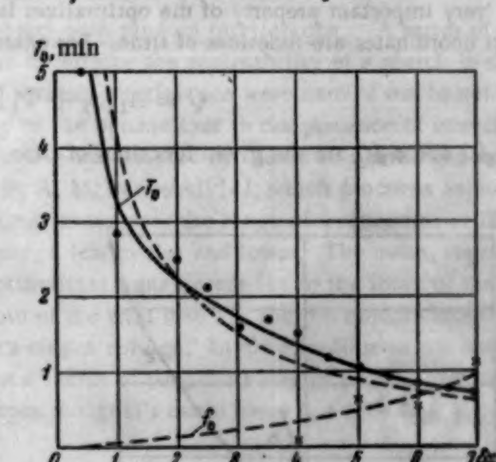


Fig. 10

The broken lines are the experimental curves calculated from the formulas [2]:

$$T_0 = \frac{\ln (x_{0\ max}/\delta)}{\ln [1/(1 - 2aA\delta x)]}, \quad f_0 = 0.25\,\delta x^2 (A_1 + A_2),$$

where x_{0max} is the initial value of the maximum coordinate ($x_{0max} = 60$ v), δ is the magnitude of the zone

around the minimum, in which the transient process of the search is considered to be finished ($\delta = 6$ v), Δt is the time taken to make one step ($\Delta t = 10.5$ sec), $\underline{a}$ is the proportionality factor between the size of the increments in the coordinates and the magnitude of the partial derivative ($a = 2$), $A_1 \approx A_2 = A$ is the coefficient of the squared term in the object's equation, $A = 0.0215$, δx is the size of the trial step.

The figure shows good agreement of the theoretical and experimental curves $T_0(k)$, if one takes into account the fact that the object's characteristic is not an exact quadratic.

Comparison of the search diagrams shows that the steepest descent method leads to a more rapid approach to the minimum than does the gradient method. However, close to the minimum, the search time with the chosen method of locating the minimum is practically the same as the search time with the gradient method. At the same time, the accuracy in determining the value of Q_{pre} in the case where the gradient method is used, is considerably greater than in the case of the steepest descent method, because of the existence of an operating threshold in the minimum-locating circuits.

Thus, it is expedient, in setting up the first initial search step, to utilize the steepest descent method, while, when in the close vicinity of the minimum, which can be judged from the value of ξ, the gradient method should be applied.

Object with variable parameters. Experiments with an object with variable parameters were conducted with the gradient method of searching, since, as was already pointed out, it is necessary to use this method of searching in the vicinity of the minimum.

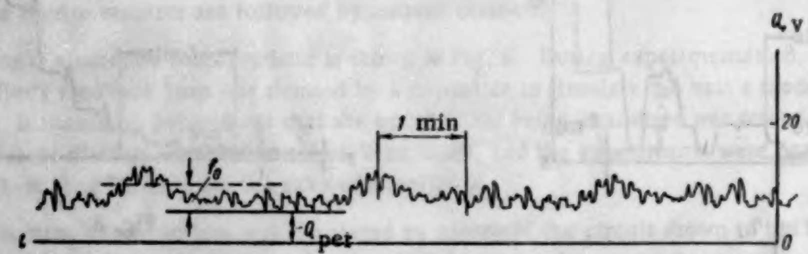


Fig. 11

A very important property of the optimalizer is its ability to track the moving extremum point in the case where its coordinates are functions of time. For example, let this function have the form:

$$Q = A_1[x_1 + \Phi_1(t)]^2 + A_2[x_2 + \Phi_2(t)]^2,$$

where $\Phi_1(t)$ and $\Phi_2(t)$ are the given functions of time.

It can be shown that there exists a critical value of the derivatives $\Phi'_1(t)$ and $\Phi'_2(t)$ when the optimalizer ceases to track the minimum effectively, and the searching becomes disorganized, since the system begins to make false steps.

In fact, between the time when the function $Q(x_{10}, x_{20})$ is memorized and the instant when the quantity ΔQ is memorized, there elapses a time Δt. If, during this time, $\Phi_1(t)$ and $\Phi_2(t)$ change sufficiently, then the effective value of the increment ΔQ_{eff}, which is memorized in terms of the values of the partial derivatives, can be larger or smaller than the true value of ΔQ, which is obtained when the object's parameters are independent of time. A critical speed will be attained when the magnitudes and signs of the derivatives $\Phi'_1(t)$ and $\Phi'_2(t)$ are such that ΔQ_{eff}, when decreasing to zero, changes sign. This is obtained when [2]

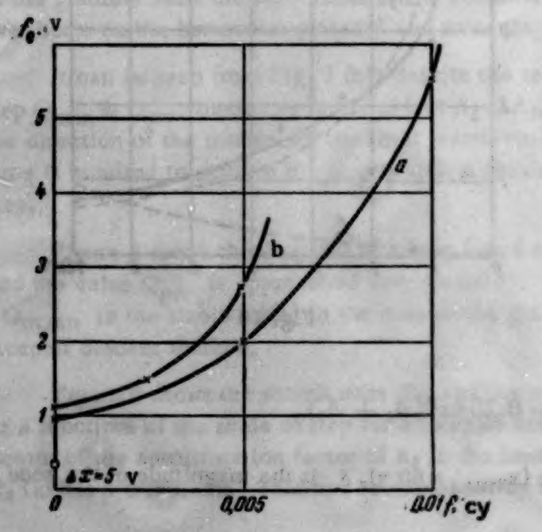


Fig. 12

$$\Phi_i'(t) = \delta x / \Delta t \quad (i = 1, 2),$$

where δx is the size of the trial step, Δt is the time interval between the memorizing of the quantities Q (x_{10}, x_{20}) and ΔQ_1.

As a result of this situation, when the coordinates of the minimum oscillate with a frequency Ω ($\Phi_1(t)$, $\Phi_2(t)$ are sinusoidal functions of time with amplitudes A_{x_1} and A_{x_2}) the output, when in a stable state, also contains harmonic components with frequency Ω, 2Ω, 3Ω, etc. A plot of the output value of Q, obtained when $\Phi_1(t) = 11 \sin \frac{2\pi}{180} t$, and $\Phi_2(t) = 9 \sin \frac{2\pi}{202.5} t$, is shown in Fig. 11.

Figure 12 shows a plot of f_0 as a function of frequency of oscillation, for constant amplitude of oscillation, for the case of oscillations along one coordinate only (curve a), and for the case of simultaneous oscillations along both coordinates (curve b) with $\delta x = 15$ v. The mean deviation was determined as the ordinate of a rectangle with its area equivalent to the output oscillations. In case b, a more rapid growth of f_0 occurs; in both cases, the rate of growth of f_0 increases as the critical speed is approached during a certain part of the period.

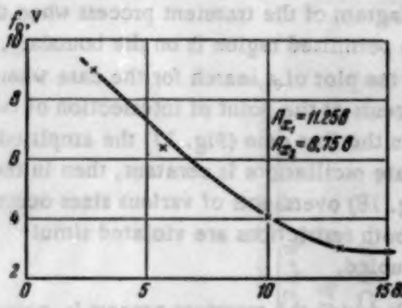


Fig. 13

One way of increasing the critical speed is to increase the trial step δx. Figure 13 shows a plot of f_0 as a function of the size of step δx when the object swings along both coordinates, and the scale of step is kept constant. From the graph, it is evident that when δx is decreased f_0 increases rapidly as the critical speed is approached. The relationship between f_0, in a stable state, and the size of the step's scale is interesting in the case of sinusoidal oscillations of the minimum, with δx = const. The graph in Fig. 14 shows that in this case, there is an optimum scale of the step, for which the quantity f_0 is a minimum. In fact, when the scale is small, the system is unable to follow the oscillating minimum, while with a very large step the system overshoots the minimum.

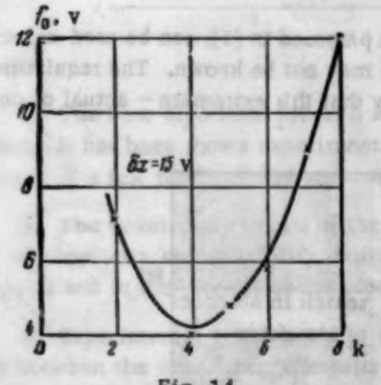


Fig. 14

An object with random interference. A series of experiments to determine the realizability of a search in the presence of random interference were carried out to determine the stability of the optimalizer in the presence of interference. The interference was created by means of a noise generator, developed by A. M. Petrovskii [4], which produces approximately "white" noise over the range of frequencies of interest to us, namely, a few cycles and lower. The noise, together with the optimalizer signal, were fed to the input of the object, so that the signal at the object's input appeared as the sum of the trial step δx and the random interference, the scale of which could be altered by changing the generator's output voltage. In the experiments the noise signal was varied from zero to rms values exceeding, by about a factor of two, the value of δx. Comparisons were made by recording on the potentiometer EPP-09 the composite signal's coordinates $x_i + \delta x + f_{\text{ran. noise}}$ ($i = 1, 2$).

Figure 15 shows the dependence of f_0 and T_0 on the output voltage of the noise generator. As can be seen from the graph, it can be assumed, as a rough approximation, that f_0 is proportional to the noise level. At the same time, T_0 is relatively independent of the noise level. This situation is explained by the appreciable suppression of interference due to the persistence of the memory loops. These, being high-pass filters, neutralize noise signals having a zero constant component, but are very sensitive to the quantity δx, i.e., to the constant component of the composite signal. For objects discrete in time this type of interference suppression is impossible because the interference comes through in the variation of the quantity ΔQ, which appears as a constant quantity at discrete instants of time. When operating with discrete objects, it is necessary to repeat the supplied δx and to

average the resulting series of values of ΔQ. Figure 16 shows a graph of the search $x_1 = f(t)$ in the absence of noise and in the presence of very strong noises.

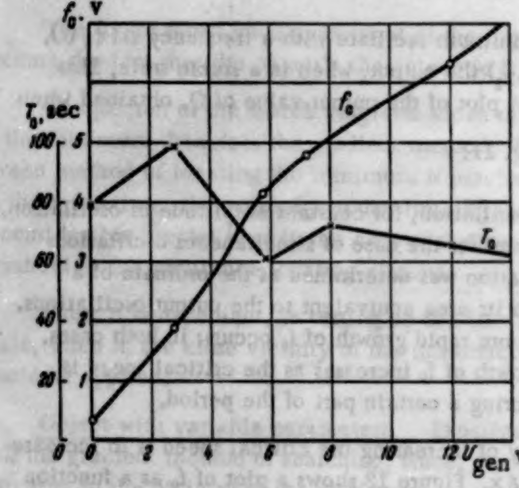


Fig. 15

Object with added restrictions. In a hunt, in the presence of added restrictions, two situations can arise: a) the minimum is in a "permitted" region, b) the minimum is on the boundary of a forbidden region.

In case "a" the search proceeds in the same way as in the absence of restrictions, except, maybe, in the case where part of the path is along the boundary of the region. The location of a minimum situated on the boundary is the more interesting case. It was indicated in [1] that, in this case, the stabilized process oscillates in the vicinity of that point where the gradients of Q and H_0 are directed exactly toward each other. Figure 17 is a diagram of the transient process when the minimum in the permitted region is on the boundary. Figure 18 shows the plot of a search for the case when the minimum occurs at the point of intersection of two boundaries. If in the first case (Fig. 17) the amplitude of the steady-state oscillations is constant, then in the second case (Fig. 18) overshoots of various sizes occur. They occur when the operating point is pushed out of the region where both restrictions are violated simultaneously and the absolute value of the gradient H_0 is approximately doubled.

It is apparent from the diagrams that the descent along a boundary during the transient process is, naturally, slower than an unrestricted search, the reduction in search speed being by approximately 1.5-2.5 times.

SUMMARY

1. The described instrument, designed on the basis of the principle proposed in [1], can be used to locate the extremum point of a function of several variables, the shape of which may not be known. The requirements imposed on the object the extremum of which is being located are merely that this extremum – actual or conditional (restrictions taken into account) – exists.

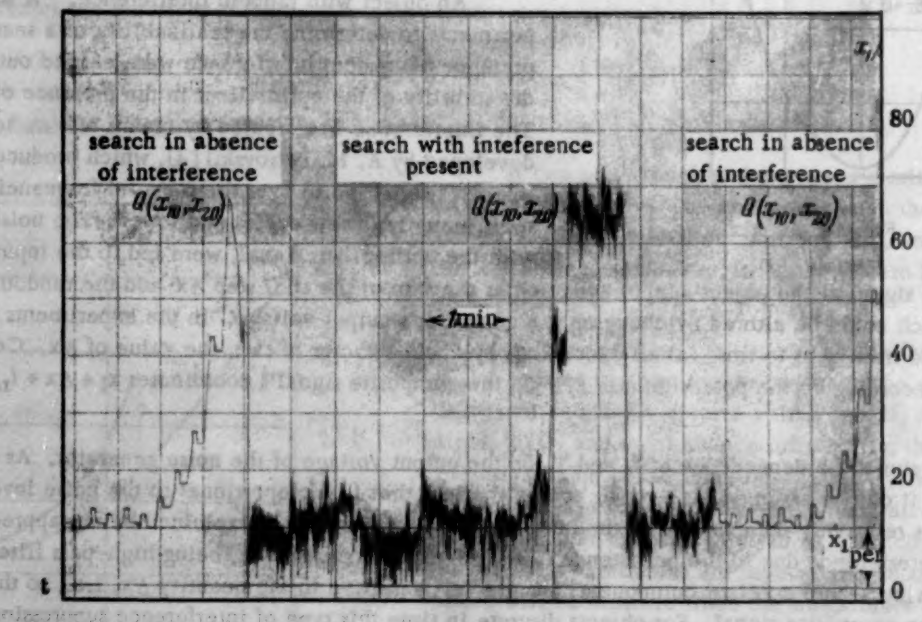


Fig. 16

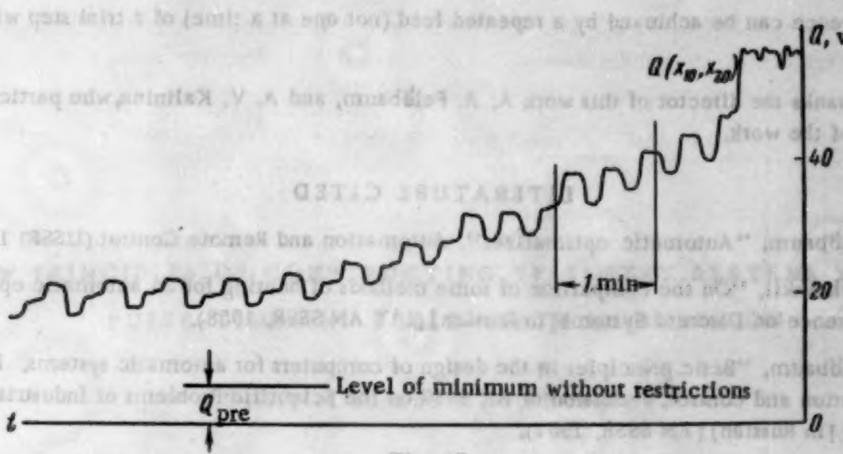


Fig. 17

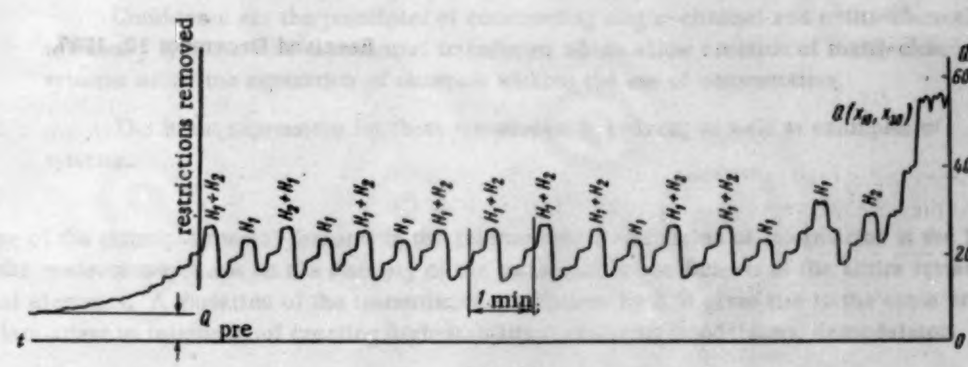


Fig. 18

2. The best algorithm for such a device is a combination of the method of steepest descent and the gradient method. It has been shown experimentally that it is possible, in the determination of $Q_{min}(Q_{pre})$, to attain an accuracy of a few tenths of one percent with respect to the maximum possible value of Q.

3. The developed circuits of the elements of the equipment turned out to be satisfactory from the point of view of simplicity and reliability, but there is still room for further refinements both in the over-all structure of the apparatus and in the circuits of the separate units.

4. Experimental verification of theoretical relationships (where this was possible) showed satisfactory agreement between the experimental results and those obtained from design formulas derived in [2].

5. The experimental study of the equipment's characteristics showed that the parameters of the optimalizer which have the greatest effect on the hunting process are the amplification factor of the input differentiator and the value of the trial step δx, as well as their product, proportional to the scale of the work step.

Whether the object's parameters are variable or constant, there is a definite, optimum scale of step, which gives the minimum value of the mean deviation in the stable state. For each value of δx there is a maximum permissible speed of motion of the extremum point. The direction of the critical velocity depends on the direction of the trial step δx, and the permissible volocity increases with increase in δx. It should be pointed out, however, that a small change in the method of determining the gradient, namely the allowance for the time derivative $Q'(t)$, permits a considerable increase in the permissible velocity of the minimum's variations.

A study of stability in the presence of interference showed that, when objects which are continuous in time are involved, the device copes with interference quite successfully. Here the signal-to-noise ratio reached 1:2 at least. An increase in the noise level results in an increase in the search time and the amplitude of the stable oscillations, but does not interfere with the operation of the equipment. An increase in the stability in the

presence of interference can be achieved by a repeated feed (not one at a time) of a trial step with a subsequent averaging of ΔQ.

The author thanks the director of this work A. A. Fel'dbaum, and A. V. Kalinina, who participated in the experimental part of the work.

LITERATURE CITED

[1] A. A. Fel'dbaum, "Automatic optimalizer", Automation and Remote Control (USSR) 19, 8 (1958).*

[2] R. I. Stakhovskii, "On the comparison of some methods of hunting for an automatic optimalizer," Report at the Conference on Discrete Systems [in Russian] (IAT AN SSSR, 1958).

[3] A. A. Fel'dbaum, "Basic principles in the design of computers for automatic systems. Basic problems of automatic regulation and control, " Session of AN SSSR on the Scientific Problems of Industrial Automation, 15-20 October,1956 [in Russian] (AN SSSR, 1957).

[4] A. M. Petrovskii, "Ultrasonic, fluctuating noise generator," Automation and Remote Control (USSR) 14, 4, (1953).

Received December 12, 1957.

* Present issue; translation page 718.

NEW PRINCIPLES OF CONSTRUCTING TELEMETRY SYSTEMS WITH PULSE-TIME AND PULSE-WIDTH MODULATION

V. A. Il'in and A. I. Novikov

(Moscow)

Considered are the principles of constructing single-channel and multi-channel telemetry systems with exponential transducers which allow creation of multi-channel systems with time separation of channels without the use of commutators.

The basic expressions for these transducers are given, as well as examples of systems.

One of the principal special features in the transmission of information is the linear dependence of the measurement error on the stability of the transmission coefficients of the entire system and of its individual elements. A variation of the transmission coefficient by δ % gives rise to the same error. Therefore, the problem arises in telemetry of creating high-stability transducers (modulators, demodulators, etc.).

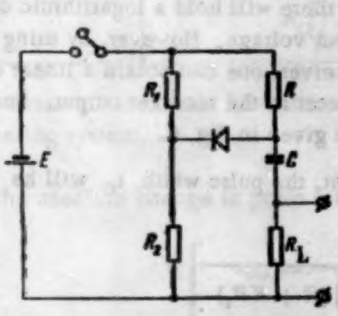


Fig. 1. Circuit of an exponential pulse-width transducer.

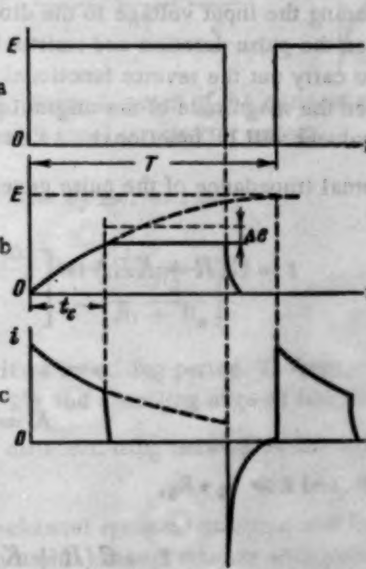


Fig. 2. Time diagrams of the voltage and current in the transducer.

As a result of work done on this problem at the Institute of Automation and Remote Control of the AN SSSR, new, simple, high-stability pulse-time and pulse-width transducers, called exponential, were suggested and developed. This paper describes such transducers and their underlying theory, and considers examples of telemetering systems.

The exponential pulse-width transducer consists of a bridge (Fig. 1), in one arm of which are connected resistors R_1 and R_2, and in the other arm of which are connected a resistor and a capacitor or a resistor and an inductance. A diode forms the bridge's diagonal. In the circuit of Fig. 1, $R_L \ll R$ and the back resistance of the diode, $R_{d0} \gg R$. If a rectangular voltage (Fig. 2a) is impressed on the bridge, the voltage on condenser C varies as shown in Fig. 2b.

At the moment of time t_c, the voltage on the condenser reaches the voltage on resistor R_2 and the diode begins to pass current. Thus, from the moment t_c until the end of the pulse, the voltage on the condenser remains constant, since the direct resistance of the diode, $R_{dd} \ll R$, $R \gg R_2$.

Figure 2c shows the variation of current through the condenser or, alternatively, the voltage on load resistor R_L.

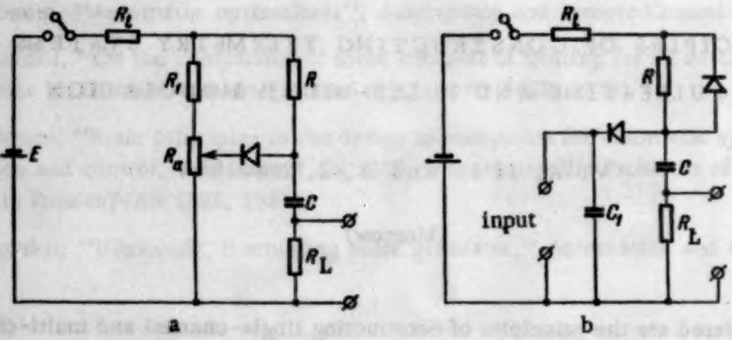


Fig. 3. Circuits for exponential transducers.

The width t_c of the pulse at the output load does not depend, with this circuit, on variations in the source voltage:

$$t_c = RC \ln \frac{R_1 + R_2}{R_1}.$$

Choosing as our data unit either resistor R or capacitor C, we obtain a linear exponential transducer, in which the dependence between the pulse width t_c and the data unit resistor R or capacitor C will be linear.*

In many cases it is expedient to use as a data unit a low-ohmic resistor R_a (Fig. 3a) or to control the transducer by feeding the input voltage to the diode (Fig. 3b). With this there will hold a logarithmic dependence between the pulse duration and resistor R_a or the transducer's input voltage. However, by using a similar circuit to carry out the reverse functional transformation at the receiver, one can obtain a linear dependence between the magnitude of the original parameter and the parameter at the receiver output. Such an exponential transducer will be functional. An example of such a system is given in Fig. 6.

If the internal impedance of the pulse generator is taken into account, the pulse width, t_c will be

$$t = C(R + KR_i) \ln \left[1 + \frac{\Delta R_a}{R_a} \frac{R}{\left(1 + \frac{R_1}{R_a} - \frac{\Delta R_a}{R_a}\right)(R + KR_i)} \right],$$

where

$$K = \frac{R_1 + R_a}{R_i + R_1 + R_a}.$$

For $R_i \ll R$ and $R \gg R_1 + R_a$,

$$t_c = C(R + KR_i) \ln \left[\frac{1}{1 - \frac{\Delta R_a}{R_1 + R_a}} \right].$$

If the internal impedance varies by ΔR_i, the relative error is

$$\delta = \frac{K \Delta R_i}{R + KR_i}.$$

* The idea of a linear exponential transducer was suggested by V. A. Il'in.

The absolute change in pulse width for $R_1 \ll R$ and $R \gg R_1 + R_a$ will be

$$\alpha_{R_1} \approx CK\Delta R_1 \ln\left[\frac{1}{1-\dfrac{\Delta R_a}{R_1+R_a}}\right].$$

The characteristics of actual diodes differ from the idealized piece-wise linear curves, so that the diodes begin conducting current earlier than the moment of time t_c. This leads to a broadening of the trailing edge of the pulses and to a dependence of the pulse width on the operating threshold Δe of the diode and the shaping circuit (Fig. 2b).

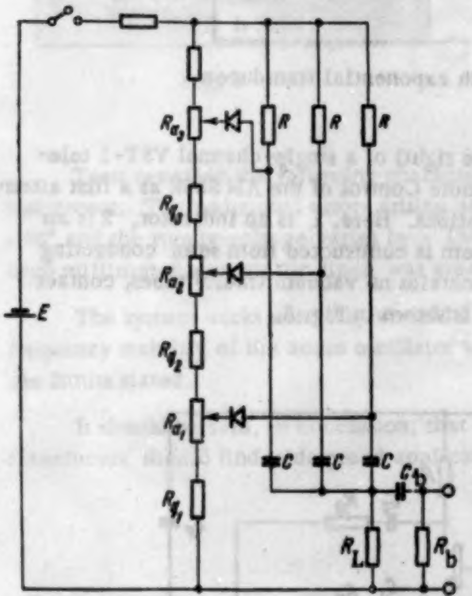


Fig. 4. Transmitter circuit for a three-channel telemetering system.

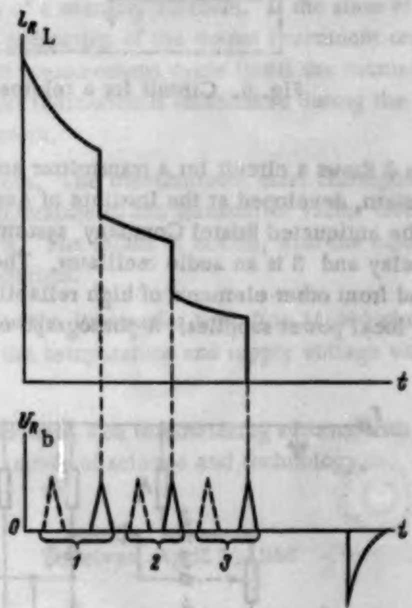


Fig. 5. Time diagram of the currents and voltages in a three-channel transmitter system.

The absolute change in pulse length when the threshold varies by Δe, will be

$$\alpha_{\Delta e} = -C(R + KR_1)\ln\left[1 - \frac{\Delta e}{E}\,\frac{1}{1-\dfrac{\Delta R_a}{R_1+R_a}}\right].$$

With actual elements, a transducer can be constructed with a repetition period T from several tens of microseconds up to tens of seconds with an error not exceeding 1% and a trailing edge of less than 1%. *

Pulse-time modulation is implemented by introducing a differentiating network at the transducer output and pulse-shaping stages in the transmitter.

Besides the well-known methods of implementing multi-channel systems with time and frequency separation, there is a good deal of interest in methods of constructing multi-channel systems with pulse-time modulation and time separation of the channels without the use of a commutator. Figure 4 shows a simplified transmitter circuit for such a three-channel system. Here, R_{a_1}, R_{a_2} and R_{a_3} are potentiometer data units of the first, second and third channels and resistors R_{q_1}, R_{q_2} and R_{q_3} are used for creating the shielding time intervals between the channels. The differentiating network R_bC_b is connected to the circuit output.

* The idea of the reverse functional transducer was suggested by A. I. Novikov.

We note that if, at the input of the transducer channels, a voltage is applied, then the circuit is set up in accordance with Fig. 3 b.

Figure 5 is illustrative of the maximum time deviations in the channels and of the shielding intervals between channels.

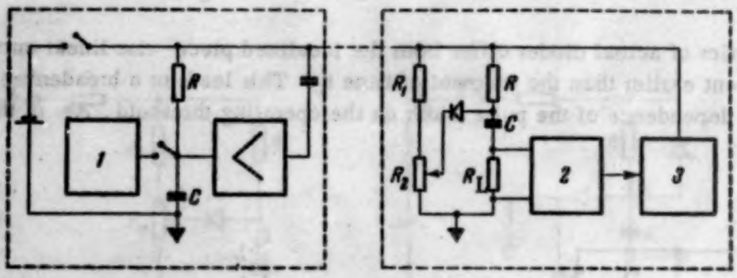


Fig. 6. Circuit for a telemetering system with exponential transducers.

Figure 6 shows a circuit for a transmitter and receiver (on the right) of a single-channel VST-1 telemetering system, developed at the Institute of Automation and Remote Control of the AN SSSR as a first attempt to replace the antiquated Bristol Company system and its modifications. Here, 1 is an indicator, 2 is an electronic relay and 3 is an audio oscillator. The transmitter system is constructed from semi-conducting elements and from other elements of high reliability (Fig. 7). It contains no vacuum (radio) tubes, contact elements or local power supplies. A photograph of the transmitter is shown in Fig. 8.

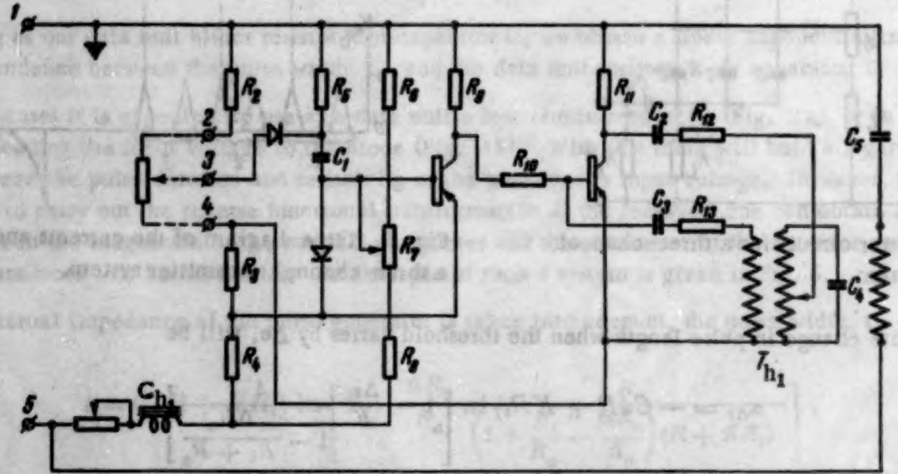


Fig. 7. Transmitter circuit of a telemetering system.

The dc pulses generated in the receiver have a width of 1.5 seconds and a repetition period of 2 seconds (Fig. 2a). They are fed to the transmitter system by line. The working pulse voltage at the transmitter input varies from 25 volts to 60 volts, with an input impedance of about 1200 ohms. The system was designed for standard potentiometric data units with impedances from 300 ohms to 2000 ohms.

The electronic relay, composed of germanium triodes, is connected to the output of the exponential transducer. The output triode of the electronic relay is simultaneously an audio oscillator, with a constant frequency chosen within the limits of 200 cycles to 2000 cycles.

At the instant t_c (Fig. 2), the electronic relay operates and switches in the audio oscillator until the end of the initial pulse.

At the receiver input is a resonant circuit, tuned to the frequency of the audio oscillator. The voltage in the resonant circuit is amplified, and causes the relay to operate at the instant t_c. The relay switches condenser C into the output stage for a brief period, and the output stage then "remembers" the condenser voltage until the following cycle.

Fig. 8

The RC networks in the transmitter and receiver have identical time constants. This guarantees a linear relationship between the magnitude of the measured quantity and the indication of the output instrument (an inverse functional transformation).

The distinguishing characteristic feature of the receiver is the inclusion of a memory element. If the state of the data unit changes, the indication of the output instrument remains constant until the next measurement cycle (until the moment t_c), after which the next indication is established during the time constant of the instrument.

Tests revealed the following characteristics of the system. The measurement error corresponds to a class I instrument. The additional errors arising as the ambient temperature of the transmitter varies from − 50° to + 50° and the supply voltage varies by ± 15% are less than 1%. The radius of action, with the equivalent of four-millimeter bi-metallic lines, was greater than 500 kilometers.

The system works normally when the series resistance in the line varies from 0 to 10,000 ohms. The frequency stability of the audio oscillator was ± 1.5% when the temperature and supply voltage varied within the limits stated.

It should be said, in conclusion, that exponential transducers, and telemetering systems with exponential transducers, should find widespread application in various branches of science and technology.

Received April 1, 1958

NOISE STABILITY OF TELEMETRY SIGNAL TRANSMISSION ON CHANNELS WITH NOISE FLUCTUATIONS

V. A. Kashirin and G. A. Shastova

(Moscow)

The potential noise stability of transmission parameters is determined under amplitude, frequency, pulse-frequency, pulse-time, pulse-width and pulse-code modulation. Generalizations of the transmission parameters are introduced. It is shown that in some range of the signal-to-noise ratio, the very best noise stability is attained with frequency modulation.

INTRODUCTION

We consider a transmission parameter λ varying within the limits ± 1 in a random way, but such that, with a sufficiently high accuracy, it can be represented by its values at moments of time which are separated by a fixed time interval T. The interval T is related to the highest frequency F_m of parameter variation by the well-known relationship

$$T = \frac{1}{2F_m}. \tag{1}$$

We shall assume that all parameter values in the interval ± 1 are equally probable.

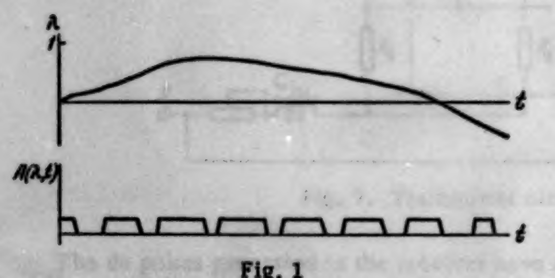


Fig. 1

Let the parameter be transmitted along the channel by means of a signal $A(\lambda, t)$ which, in the general case, is a function both of time and of the transmitted parameter. The function $A(\lambda, t)$ depends on the type of modulation. Figure 1 shows the time variation of the parameter λ and the function $A(\lambda, t)$ for pulse-width modulation.

When the parameter λ_0 is transmitted along a channel with weak noise fluctuations, the specific intensity of which equals σ volts/cycle, errors arise, the relative value (over the entire range) of which may be determined, for an ideal receiver, from the formula

$$\delta^2(\lambda_0) = \frac{\sigma^2}{8\int\limits_0^T \left[\frac{\partial A(\lambda, t)}{\partial \lambda}\right]^2_{\lambda=\lambda_0} dt} \tag{2}$$

Formula (2) was obtained from the corresponding expression in [1] for the absolute value of the mean square error by dividing the error by the magnitude of the range of variation of the transmitted parameter and by changing the limits of integration. The integration of the square of the derivative $\partial A(\lambda, t)/\partial\lambda$ is carried out in the present work over the limits 0 to T, while in [1] the integration was over the segment from $-T/2$ to

T/2. Such a change of limits is completely admissible and is not essential from the point of view of the theory of potential noise stability, although it does influence the resulting values of the mean square errors for frequency, time and width modulation.

The mean square error over all values of λ is found from the usual integration

$$\delta^2 = \frac{1}{2}\int_{-1}^{+1} \delta^2(\lambda_0)\, d\lambda_0. \tag{3}$$

By means of the formulas given, we now determine the potential noise stability of various forms of modulation, with the condition that the dynamic range of the signal $A(\lambda, t)$ is limited, i.e., the signal can vary within the limits $-U_m$ to U_m or 0 to $2U_m$.

Amplitude Modulation (AM)

With amplitude modulation, the signal can be written in the form

$$A(\lambda, t) = U_0(1 + m\lambda)\cos\omega_0 t,$$

where U_0 is the mean value of the signal amplitude, m is the index of amplitude modulation and ω_0 is the signal carrier frequency.

For finding the error, we define the integral:

$$I_{AM} = \int_0^T \left[\frac{\partial A(\lambda, t)}{\partial\lambda}\right]^2 dt = U_0^2 m^2 \int_0^T \cos^2\omega_0 t\, dt.$$

$$I_{AM} = {}^1/_2\, U_0^2 m^2 T \quad \text{for} \quad T \gg 2\pi/\omega_0$$

By substituting the value of I_{AM} in (2), we obtain the formula for the error with AM:

$$\delta_{AM}^2 = \frac{\sigma^2}{4U_0^2 m^2 T}.$$

For convenience in comparing the various transmission methods, we introduce the generalized parameter ρ, equal to the ratio of the maximum signal voltage to the effective noise voltage in the frequency band taken by the parameter λ:

$$\rho = \frac{U_m}{\sigma\sqrt{F_m}} = \frac{U_m\sqrt{2T}}{\sigma} = \left(\frac{U_s}{U_n}\right)_{\Delta F = F_m}. \tag{4}$$

By taking into account that with AM, $U_m = U_0(1 + m)$, we get $\delta_{AM} = \dfrac{1+m}{\rho m\sqrt{2}}$.

The minimum error is for $m = 1$,

$$\delta_{AM\,\min} = \sqrt{2}/\rho. \tag{5}$$

Thus, for amplitude modulation, the error is completely determined by the parameter ρ and increases in inverse proportion to it. Formula (5) remains valid for any noise level [1].

Frequency Modulation (FM)

Using frequency modulation, with carrier frequency f_0 and deviation frequency f_d, the signal can be presented in the form

$$A(\lambda, t) = U_m \sin 2\pi(f_0 + f_d\lambda)t.$$

The integral sought is

$$I_{FM} = U_m^2 (2\pi)^2 f_d^2 \int_0^T t^2 \cos^2 2\pi (f_0 + f_d \lambda)\, t dt.$$

If $T \gg 2\pi / f_0$, then $I_{FM} = \frac{2}{3}\pi^2 U_m^2 f_d^2 T^3$, and the error is

$$\delta_{FM} = \frac{0.14\,\sigma}{U_m f_d T^{3/2}}.$$

We now introduce the parameter γ, equal to the ratio of one-half the band taken by the signal to the frequency band of parameter λ:

$$\gamma = \frac{\Delta f}{2F_m}. \tag{6}$$

With frequency modulation, $\frac{\Delta f}{2} = f_d + F_m$ and

$$\gamma_{FM} = \frac{f_d}{F_m} + 1. \tag{7}$$

By taking (1), (4) and (7) into account, we obtain the expression for the mean square error using frequency modulation:

$$\delta_{FM} = \frac{0.4}{\rho(\gamma - 1)}. \tag{8}$$

In contradistinction to AM, the error using FM depends not only on parameter ρ, but also on the quantity γ, characterizing the degree of band expansion needed by the signal in comparison with AM. As is obvious from (8), for $\gamma \gg 1$, the error, with weak noise, increases in inverse proportion to the first power of the broadening coefficient of the frequency band.

The use of secondary AM impairs the noise stability [2].

With this, the mean square error is

$$\delta_{FM-AM} = \frac{2.2}{\rho(\gamma - 1)}.$$

Pulse-Frequency Modulation (PFM)

A. Pulse-frequency modulation with constant pulse width (PFMI).

With PFMI, the frequency of the sequential pulses $f_i(\lambda)$ which have a constant form and amplitude, varies linearly as a function of the magnitude of the measured parameter, λ:

$$f_i(\lambda) = f_i + f_d \lambda,$$

where f_i is the mean frequency of the pulses of the sequence and f_d is the deviation frequency.

The signal can be written in the form of a sum

$$A(\lambda, t) = \sum_{k=0}^{\infty} F(t - t_k),$$

where $F(t)$ is a function describing the envelopes of the individual pulses and t_k is the moment of transmission of the k'th pulse,

$$t_k = \frac{k}{f_i + f_d \lambda} \quad (k = 1, 2, \ldots).$$

The derivative $\partial A(\lambda, t)/\partial\lambda$ will equal

$$\frac{\partial A(\lambda, t)}{\partial\lambda} = \sum_{k=0}^{\infty}\frac{\partial F(t-t_k)}{\partial\lambda}.$$

In order to compute the derivatives $\partial F(t-t_k)/\partial\lambda$, we introduce a change of variables:

$$t_k^{*} = t - \frac{k}{f_i + f_d\lambda}.$$

Then,

$$I_{PFMI} = \int_0^T\left[\sum_{k=0}^{\infty}\frac{\partial F(t_k^{*})}{\partial t_k^{*}}\,\frac{\partial t_k^{*}}{\partial\lambda}\right]^2 dt.$$

Since the pulses are nonoverlapping, and the derivative $\partial F(t_k)/\partial t_k \neq 0$ only during the time when the k'th pulse exists (where the magnitude of the derivative is independent of the number k), the desired integral equals

$$I_{PFMI} = \sum_{k=0}^{\infty}\frac{k^2 f_d^2}{(f_i + f_d\lambda)^4}\int_{t_k-\frac{\tau}{2}}^{t_k+\frac{\tau}{2}}\left[\frac{\partial F(t_k^{*})}{\partial t_k^{*}}\right]^2 dt. \tag{9}$$

We now consider transmission by means of triangular pulses of amplitude $2U_m$ and width τ.

In this case

$$\int_{t_k-\frac{\tau}{2}}^{t_k+\frac{\tau}{2}}\left[\frac{\partial F(t_k^{*})}{\partial t_k^{*}}\right]^2 dt = \int_{t_k-\frac{\tau}{2}}^{t_k+\frac{\tau}{2}}\left(\frac{4U_m}{\tau}\right)^2 dt = \frac{16U_m^2}{\tau}. \tag{10}$$

By the use of (9), (10) and (2), we find the magnitude of the squared error:

$$\delta^2(\lambda) = \frac{3\sigma^2\tau(f_i + f_d\lambda)^3}{128\,U_m^2 f_d^2\,T\,(L^2 + 1.5\,L + 0.5)},$$

where $L = T(f_i + f_d\lambda)$ is the number of pulses appearing during time T when the given parameter λ is transmitted.

As is obvious from the expression obtained, the magnitude of the squared error depends on the values assumed by the parameter λ. By remembering that all values of λ within the limits from -1 to $+1$ are equiprobable, we determine, from Formula (3), the mean square relative error, averaged over all values of λ:

$$\begin{aligned}\delta^2_{PFMI} = \frac{3\sigma^2\tau}{256\,U_m^2 f_d^2\,T}\Bigg\{&\frac{x^2}{2T^2} - \frac{1.5\,x}{T^3} + \frac{0.87}{T^4}\ln\left[0.5 + 1.5\,Tx + T^2x^2\right] + \\ &+ \frac{1.12}{T^4}\ln\left[\frac{1.5 + 2Tx - 0.5}{1.5 + 2Tx + 0.5}\right] - \frac{y^2}{2T^2} + \frac{1.5\,y}{T^3} - \frac{0.87}{T^4}\ln\left[0.5 + 1.5\,Ty + T^2y^2\right] - \\ &- \frac{1.12}{T^4}\ln\left[\frac{1.5 + 2Ty - 0.5}{1.5 + 2Ty + 0.5}\right]\Bigg\},\end{aligned} \tag{11}$$

where $x = f_i + f_d$ and $y = f_i - f_d$.

If $L \gg 1$ (or $f_i \gg f_d$), i.e., if during time T for any λ, a large number of pulses is transmitted and $\tau = \frac{1}{f_m} \approx \frac{1}{f_i}$, then the expression for the squared error simplifies to:

$$\delta^2_{PFMI} = \frac{3\sigma^2}{128\,f_d^2\,U_m^2\,T^3}. \tag{12}$$

Taking into account that $\Delta f_{PFM} = 2f_d$, and making use of (4) and (6), we obtain an expression for the error in generalized parameters:

$$\delta_{PFMI} = \frac{0.4}{\rho\gamma}. \tag{13}$$

On the basis of (11) it may easily be shown that the error is significantly decreased if the deviation frequency is increased. With given frequency limits f_1 and f_m, the maximum value of f_d is attained for

$$f_i = \frac{f_m + f_1}{2}. \tag{14}$$

It will be assumed in the sequel that Condition (14) always holds. In the case when the frequencies f_m and f_1 are not given, but only their difference $\Delta f = f_m - f_1$, is given, the problem arises of choosing the working portion of the frequency spectrum. By using (12), one can assure oneself that the derivative $\partial\delta / \partial f_1$ is positive. Consequently, from the point of view of noise stability, it is advantageous to decrease f_1, given the band Δf. If the highest frequency of variation of the transmitted parameter is F_m, then the least value of f_1 equals $2F_m$.

Formula (12) is valid, however, only for $f_1, f_i \gg f_d$.

To investigate the dependence of δ on f_1 for small f_1, we return to general Formula (11). By taking $f_1 = k(2F_m)$ and $f_d \gg f_i$, we find that even for the low-frequency portion of the spectrum, $\partial\delta / \partial k > 0$, i.e., the error also falls when the lowest frequency of the range is decreased.

Thus, the minimum error occurs when

$$f_1 = 2F_m. \tag{15}$$

By substituting (15) in (11), and using generalized parameters, we obtain an expression for the minimum error (in the case when $f_1 = 2F_m$):

$$\delta_{PFMI} = \frac{0.3}{\rho\gamma}\left[1 - \frac{3}{\gamma} + \frac{9.2}{\gamma^2} + \frac{2.2}{\gamma^3}\ln\left(\frac{1+2\gamma}{2+2\gamma}\right)\right]^{1/2},$$

for $\gamma > 2$, the error, with an accuracy of 10%, is determined from the formula

$$\delta_{PFMI} = \frac{0.3}{\rho\gamma}\left(1 - \frac{3}{\gamma} + \frac{9.2}{\gamma^2}\right)^{1/2}. \tag{16}$$

As is clear from (16) and (13), a shift of the signal frequency band to the low-frequency region decreases the error. For large γ, the error is decreased by a factor of approximately $\sqrt{2}$. In practice, however, it is not always possible to use just the low-frequency portion of the spectrum.

Previously we considered transmission by triangular video pulses. It may be shown, by analogous methods, that with the same dynamic range, the error for PFMI-AM, i.e., for transmission by radio pulses with a triangular envelope of amplitude U_m, for $\gamma > 2$, will be

$$\delta_{PFMI-AM} = \frac{1.7}{\rho\gamma}\left(1 - \frac{3}{\gamma} + \frac{9.2}{\gamma^2}\right)^{\frac{1}{2}}.$$

B. Pulse-frequency modulation with constant porosity (PFMII).

With PFMII, the width of the pulses and the frequency with which they follow one another, vary proportionally to the magnitude of the measured parameter in such a manner that the porosity remains constant and equal to two.

The signal is

$$A(\lambda, t) = \sum_{k=0}^{\infty}[F(t - t_k), \lambda],$$

where $F(t, \lambda)$ is a function describing the individual pulses and depending on the parameter λ.

The square of the derivative $\partial A(\lambda, t)/\partial\lambda$ can be determined by the same method used for PFMI. Computations show that $\delta_{PFMI} = \delta_{PFMII}$.

Table 1 gives the formulas for computing the relative mean square error for various cases of PFM.

TABLE 1

Form of modulation	Mean-square-error for weak noise	Form of modula-tion	Mean-square-error for weak noise
AM	$\frac{\sqrt{2}}{\rho}$	PTM	$\frac{0.25}{\rho\sqrt{\gamma}}$
FM	$\frac{0.4}{\rho(\gamma-1)}$, for $\log\rho > \frac{\log\gamma+1}{1.8}$	PTM-AM	$\frac{1}{\rho\sqrt{\gamma}}$
FM -AM	$\frac{2.2}{\rho(\gamma-1)}$ for $\log\rho > \frac{\log\gamma+1}{1.8}$	PWMI	$\frac{0.35}{\rho\sqrt{\gamma}}$
PFMI $f_1 = 2F_m \ll f_g,\ \gamma > 2$	$\frac{0.3}{\rho\gamma}\left(1-\frac{3}{\gamma}+\frac{9.2}{\gamma^2}\right)^{1/2}$	PWMII	$\frac{0.49}{\rho\sqrt{\gamma}}$
PFMI $f_1,\ f_m \gg f_g$	$\frac{0.4}{\rho\gamma}$	PWMI AM	$\frac{1.4}{\rho\sqrt{\gamma}}$
PFMI-AM	$\frac{1.7}{\rho\gamma}\left(1-\frac{3}{\gamma}+\frac{9.2}{\gamma^2}\right)^{1/2}$	PWMI AM	$\frac{1.96}{\rho\sqrt{\gamma}}$
PFMII $f_1 = 2F_m \ll f_g,\ \gamma > 2$	$\frac{0.3}{\rho\gamma}\left(1-\frac{3}{\gamma}+\frac{9.2}{\gamma^2}\right)^{1/2}$	PCM	$\left[\frac{1}{12(2^n-1)^2} + A(n)V\left(\frac{\beta\rho}{\sqrt{n}}\right)\right]^{1/2}$
PFMII $f_1,\ f_m \gg f_g$	$\frac{0.4}{\rho\gamma}$	PCM $n = n_{opt}$	$10^{-0.3-0.2\beta\rho}$ $3.5 < \beta\rho < 25$

Note. The text provides the limits of applicability of these formulas.

Based on the results obtained, the following conclusions may be drawn.

1. For the same dynamic range, PFMI is equivalent in noise stability, to PFMII. In noise stability, PFMI-AM is worse than PFMI by a factor of $4\sqrt{2}$.

2. In all the cases of PFM considered, the magnitude of the error depends in the same ways on ρ and on γ, where this dependence is the same as with FM.

3. For operation in the low-frequency portion of the spectrum ($f_1 = 2F_m \ll f_d$), the noise stability of PFM is somewhat higher than that of FM.

For operation in the high-frequency portion of the spectrum $f_1, f_m \gg f_d$, PFM is equivalent, from the point of view of noise stability, to FM.

Pulse-Time and Pulse-Width Modulation (PTM and PWM).

With PTM, the signal is a sequence of pulses of constant form which are shifted in time with respect to the interval markers, these shifts being proportional to the parameter λ. Let the period of the marker points be T and for variations of the parameter λ from -1 to $+1$, let the time shift of the pulses vary from $-T/2$ to $T/2$. Denoting the pulse envelope by $U(t)$, we write the signal in the form

$$A(\lambda, t) = \sum_{k=0}^{\infty} U\left[t - \left(kT + \frac{1}{2}T\lambda\right)\right].$$

To determine the derivative $\partial U/\partial\lambda$, we carry out a change of variables $t' = t - (kT + T\lambda/2)$. Then,

$$\left|\frac{\partial A}{\partial\lambda}\right| = \left|\frac{\partial U}{\partial t'}\cdot\frac{\partial t'}{\partial\lambda}\right| = \left|\frac{\partial U}{\partial t'}\right|\frac{T}{2}.$$

For simplicity of analysis, we consider transmission by triangular pulses with amplitude $2U_m$ and width τ. At the moment of pulse occurrence,

$$\left|\frac{\partial U}{\partial t'}\right| = \frac{4U_m}{\tau} \text{ and } \left|\frac{\partial U}{\partial\lambda}\right| = \frac{4U_m}{\tau}\frac{T}{2} = \frac{2U_m T}{\tau}.$$

Since just one pulse occurs during time T, the integral of the square of the derivative $\partial A/\partial\lambda$ equals

$$I_{PTM} = \frac{4U_m^2 T^2}{\tau^2}\int_{\left(kT+\frac{T}{2}\lambda\right)-\frac{\tau}{2}}^{\left(kT+\frac{T}{2}\lambda\right)+\frac{\tau}{2}} d\tau = \frac{4U_m^2 T^2}{\tau}.$$

Substituting the magnitude of I_{PTM} in the expression for the error, we obtain

$$\delta_{PTM} = \frac{\sigma\sqrt{\tau}}{\sqrt{32}U_m T}.$$

For PTM, the parameter γ will be

$$\gamma_{PTM} = \frac{0.5}{\tau F_m}. \tag{17}$$

Expressing the error with PTM in generalized parameters, we get

$$\delta_{PTM} = \frac{0{,}25}{\rho\sqrt{\gamma}}.$$

Computations show that transmission by radio pulses, i.e., by the use of a secondary AM on the carrier frequency, will entail an error, for the same dynamic range and band, four times as great:

$$\delta_{PTM\text{-}AM} = \frac{1}{\rho\sqrt{\gamma}}.$$

Thus, exactly as for FM, the error in PTM is determined by the parameters ρ and γ. However, unlike FM, the error in PTM varies in inverse proportion to the square root of the frequency band-width taken by the signal. With PTM, the error is completely determined by the steepness and width of the pulse edge. It remains unchanged if transmission is by pulses with a trapezoidal envelope having the same edge as the triangular pulses, since the derivative $\partial A(\lambda, t)/\partial\lambda$ equals zero on the flat portion of the pulses.

It is easily shown that, with one-sided PWM the error will be larger by a factor of $\sqrt{2}$, then with PTM, if the pulses have the same steepness and width of edge:

$$\delta_{PWMI} = \frac{0{,}35}{\rho\sqrt{\gamma}}.$$

For two-sided PWM, the error will be larger by a factor of $\sqrt{2}$, than for one-sided PWM:

$$\delta_{PWMII} = \frac{0.49}{\rho\sqrt{\gamma}}.$$

Correspondingly,

$$\delta_{PWMI-AM} = \frac{1.4}{\rho\sqrt{\gamma}} \; ; \; \delta_{PWMII-AM} = \frac{1.96}{\rho\sqrt{\gamma}}.$$

Pulse-Code Modulation (PCM)

We now estimate the noise stability of parameter transmission by a simple binary code which uses all possible combinations. The errors in such transmission arise both from quantification and from the suppression of correct, or generation of incorrect, pulses. We consider errors due to misrepresentation of the code pulses by noise. Let the code consist of pulses, the values of which increase with increasing ordinal number of the pulses in accordance with the binary law: $2^0, 2^1, 2^2, \ldots, 2^{n-2}, 2^{n-1}$. We assume that the probability that one pulse is misrepresented by noise is p_1, where suppression of a correct pulse and generation of a false pulse are equiprobable. Then, the probability that just one pulse is misrepresented in a code combination consisting of n pulses equals $p(1) = np_1(1-p_1)^{n-1}$. For $np_1 \ll 1$, $p(1) = np_1$.

We assume that the probability of two misrepresentations occurring in one code combination satisfies the inequality $p(2) \ll p(1)$. We now determine the mean square error under the condition that misrepresentation is equally probable for any pulse in the code. The mean value of the squared error, if one pulse in the code combination is misrepresented, equals

$$\delta_m^2 = \frac{1}{n}\sum_{i=0}^{n-1} 2^{2i} = \frac{4^n-1}{3n}. \tag{18}$$

By taking into account that the entire parameter range, using pulse-code modulation, is transmitted by means of $(2^n - 1)$ gradations, and that the probability of one error is $p(1)$, we determine the mean square relative error:

$$\delta_m^2 = \left(\frac{4^n-1}{3n}\right)\frac{1}{(2^n-1)^2}\,p(1) = \frac{4^n-1}{3(2^n-1)^2}\,p_1 = A(n)\,p_1, \tag{19}$$

where $A(n) = \dfrac{4^n-1}{3(2^n-1)^2}$.

For $n \gg 1$, $A(n) = \dfrac{1}{3}$.

The probability that one pulse of a binary code be misrepresented is determined, as shown in [1], by its energy, the method of transmission and by the specific noise power. For any noise level,

$$p_1 = V\left(\frac{\nu}{\sqrt{2}}\frac{Q}{\sigma}\right),$$

where Q^2 is the maximum pulse energy equal, when the envelope is rectangular with pulse duration τ, to

$$Q^2 = \frac{U_m^2\tau}{2},$$

and ν is a coefficient which depends on the method of transmitting the binary digits (form of keying modulation) [3].

The function V is defined by the expression

$$V(x) = \frac{1}{2\pi}\int_x^{\infty} e^{-\frac{1}{2}x^2}\,dx.$$

For amplitude modulation, $\nu_{AMt} = 1$, for frequency modulation, $\nu_{FMt} = \sqrt{2}$ and for phase modulation, $\nu_{PMt} = 2$. We will assume that the code pulses are transmitted with porosity

$$\alpha_p = \frac{T}{n\tau}.$$

Then,

$$p_1 = V\left(\frac{v}{2\sqrt{2\alpha_p}}\frac{\rho}{\sqrt{n}}\right). \tag{20}$$

We introduce an additional parameter of code transmission,

$$\beta = \frac{v}{2\sqrt{2\alpha_p}}, \tag{21}$$

and write

$$p_1 = V\left(\frac{\beta\rho}{\sqrt{n}}\right).$$

Substituting (20) in (19), we get

$$\delta^2_m = A(n)\, V\left(\frac{\beta\rho}{\sqrt{n}}\right). \tag{22}$$

It follows from (21) and (18) that, for a given value of the generalized transmission parameter ρ, which does not depend on the method of transmission, the error due to noise in code modulation increases with an increasing number of pulses in the code. We now determine the error due to the quantification of the signal. We assume that all values of this error are equally probable in the range $-\frac{\Delta m}{2}$ to $\frac{\Delta m}{2}$, where

$$\Delta_m = \frac{1}{2^n - 1}. \tag{23}$$

Then, the probability density of the error equals $\frac{1}{\Delta m}$, and the mean square value of the error is determined from the formula:

$$\delta^2 = \int_{-\frac{\Delta_m}{2}}^{\frac{\Delta_m}{2}} \delta^2 p(\delta)\, d\delta = \frac{1}{\Delta_m}\frac{\delta^3}{3}\Bigg|_{-\frac{\Delta_m}{2}}^{\frac{\Delta_m}{2}} = \frac{\Delta^2_m}{12}.$$

By the use of (23), we obtain

$$\delta^2 = \frac{1}{12(2^n - 1)^2}.$$

As was to be expected, the error due to quantification decreases as the number of pulses in the code increases. If we assume independence between the misrepresentations due to noise and due to quantification, we obtain the total mean square error by addition:

$$\delta^2_{PCM} = \delta^2 + \delta^2_m = \frac{1}{12(2^n - 1)^2} + A(n)\, V\left(\frac{\beta\rho}{\sqrt{n}}\right). \tag{24}$$

Since the quantification error decreases with an increasing number of code pulses and the noise error increases with increasing $\underline{n}$ then, obviously, there is some optimal number of pulses in the code for which the total error is a minimum. An analytic solution of the equation, $\partial\delta^2_{PCM}/\partial n = 0$ is mathematically difficult. For the range of values, $3.5 < \beta\rho < 25$, the functional dependence, $n_{opt} = f(\beta\rho)$, was determined by computation. With an accuracy of 7%, this dependence, in the range given, can be approximated by a linear function:

$$n_{opt} = 0.5 + 0.6\beta\rho. \tag{25}$$

With an accuracy of 5%, the minimum value of the error obtained for $n = n_{opt}$ in the range of $\beta\rho$ given above, can be calculated from the formula

$$\delta_{PCM} = 10^{-0.3-0.2\beta\rho}. \tag{26}$$

On the basis of the results obtained, the conclusion may be drawn that the noise stability of PCM is completely determined by the generalized signal parameter ρ, and code transmission parameter β. The greatest value of β, and, consequently, the least error, is guaranteed by phase keying without intervals (the magnitude of β depends on the method of keying and on the code porosity). For each value of the product $\beta\rho$, there is an optimum number of pulses in the code, for which the error will be a minimum. It is obvious that, for a given length of code, the frequency band taken by the signal is proportional to the number of pulses in the code:

$$\frac{1}{2}\Delta f_{PCM} = \frac{\mu}{\tau} = \frac{\mu n \alpha_p}{T}, \tag{27}$$

where μ is a coefficient characterizing the dependence between the duration of transmission for a binary symbol and the frequency band necessary for its transmission. For PCM the parameter γ will be

$$\gamma_{PCM} = 2\mu n \alpha_p \tag{28}$$

The optimal number of pulses in the code corresponds to the optimal value of the parameter γ:

$$\gamma_{PCM(opt)} = 2\mu\alpha_p(0.5 + 0.6\beta\rho). \tag{29}$$

Figure 2 shows the dependence of the minimum total error with PCM on the magnitude of ρ for the cases of amplitude keying with porosities of $\alpha_p = 2$ and $\alpha_p = 1$, frequency modulation without intervals ($\alpha_p = 1$) and phase modulation without intervals ($\alpha_p = 1$).

Figure 3 gives the dependence of γ_{opt} on the magnitude of ρ for various methods of keying. In computing γ, it was assumed that $\mu_{AMt} = \mu_{PMt} = 1.3$ and $\mu_{FMt} = 2.6$.

Comparative Noise Stability for Various Transmission Methods

The table gives the definitive expressions for the error in the presence of weak noise. It is clear from the table that FM, PFM, PTM and PWM permit a decrease in error, as compared to AM, at the cost of broadening the frequency band used. FM and PFM are more effective than PTM and PWM since the former allow a decrease in error inversely proportional to the first power of the frequency band-width whereas PTM and PWM allow a decrease in error which is inversely proportional to the square root of the frequency band-width. These results differ somewhat from the conclusions of work [1], where it is shown that PTM and FM are equivalent from the point of view of noise stability when there is fluctuating noise. The discrepancy in results arises from different conditions of comparison.

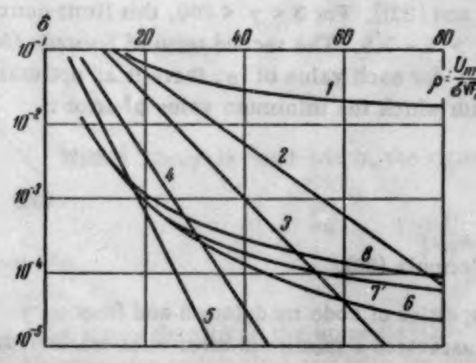


Fig. 2. Dependence of the mean square error on the parameter ρ for: 1) AM; 2) PCM-AMt, $\alpha_p = 2$; 3) PCM-AMt, $\alpha_p = 1$; 4) PCM-FMt, $\alpha_p = 1$; 5) PCM-PMt, $\alpha_p = 1$; 6) FM with optimal band; 7) FM for the band optimal (for the same ρ) for PCM-AMt for $\alpha_p = 2$; 8) FM for the band optimal for PCM-AMt for $\alpha_p = 1$.

In work [1], the modulation comparisons are made for equal signal powers, while in this work, the condition is that the dynamic ranges be equal. Systems with pulse-frequency modulation, in many cases, permit a somewhat higher degree of noise stability than systems with frequency modulation, although this increase is small. Therefore, in practice, it may be considered that, with weak noise, systems with pulse-frequency modulation are equivalent to systems with frequency modulation, from the point of view of noise stability.

We now consider in more detail systems with frequency modulation, and compare them with systems with code modulation. In systems with code modulation, as was shown above, for each ratio of signal level, noise level and speed of action, there is a certain optimal value of number of pulses in the code and corresponding to this, an optimal value of signal band-width, for which minimum error is attained. It follows from (8) that, for systems with frequency modulation, the error can be made arbitrarily small by broadening the frequency band (by increasing γ). However, this is not entirely the case. Formula (8) is valid for weak noise. When the band is broadened the noise increases and, for a definite value of γ, Formula (8) becomes invalid. In work [4], correlation methods were used to investigate the noise stability of an idealized frequency receiver, consisting of an input filter, an ideal frequency discriminator and an output filter. The mean square error with parameter transmission and reception by this receiver were determined for

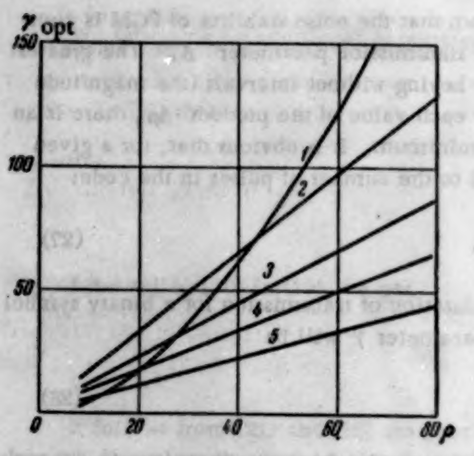


Fig. 3. Dependence of the optimal value of the parameter γ on ρ for: 1) FM; 2) PCM-FMt ($\alpha_p = 1$); 3) PCM-PMt ($\alpha_p = 1$); 4) PCM-AMt ($\alpha_p = 2$); 5) PCM-AMt ($\alpha_p = 1$).

$$\gamma > 5 \text{ and } \left(\frac{U_s}{U_n}\right)_{in} > 2.2$$

by the formula

$$\delta^2_{FM} = \frac{0.083}{\rho^2\gamma^3} + \frac{0.26}{\rho\gamma^{1/2}} e^{-\frac{\rho^2}{2\gamma}}. \qquad (30)$$

If

$$\gamma < 10^{-1+1.8 \log \rho}, \qquad (31)$$

then the second term in (30) is much less than the first term, and

$$\delta_{FM} = \frac{0.3}{\rho\gamma}. \qquad (32)$$

Expression (32) almost coincides with (8), and characterizes the error for weak noise. Inequality (31) defines the limit of applicability of the formulas for weak noise [Formulas (8) and (32)]. For $3 < \gamma < 200$, this limit corresponds to a signal-to-noise ratio at the receiver input of $(U_s/U_n)_{in} > 3 - 3.5$. The second term of Formula (30) becomes the dominant one for weak noise. It follows from (30) that for each value of ρ, there is an optimal value of the parameter γ, and a corresponding frequency band, with which the minimum value of error is obtained:

$$\gamma_{opt} = 10^{-1+1.8 \log \rho}. \qquad (33)$$

The minimum error, for $\gamma = \gamma_{opt}$, can be determined from Formula (30).

Thus, it is possible to note certain analogies between the properties of code modulation and frequency modulation. In both cases there is an optimal signal band which guarantees a minimum error when transmitting over a channel with noise fluctuations. It is interesting to compare the minimum values of error for FM and PCM. Figure 2 shows the dependence of the minimum error on the parameter ρ for PCM for various secondary modulations and for different code porosities, as well as the dependence of the minimum error on ρ for FM. It is clear from the Figure that for $7 < \rho < 18$, transmission by FM provides a lower error than transmission by PCM with the most stable, under noise, secondary keying, namely, phase keying (PCM-PMt). For $7 < \rho < 31$, frequency modulation gives a lower error than pulse-code modulation followed by frequency keying (PCM-FMt), and for $7 < \rho < 60$, FM is better in this sense than PCM followed by amplitude keying (PCM-AMt). On the basis of these results, the following general conclusion can be drawn, that for very weak noise ($\delta < 10^{-3} - 10^{-4}$), pulse-code modulation has higher noise stability than frequency modulation. In the range of error, $10^{-1} < \delta < 10^{-3} - 10^{-4}$, frequency modulation permits a lower transmission error to be obtained than does pulse-code modulation.

We now compare FM and PCM from the point of view of band-width taken by the signal. Figure 3 shows the dependence of the parameter $\gamma_{opt} = \Delta f_{opt}/2F_m$ on ρ for FM and PCM. It is clear from the figure that in the range of values of ρ for which FM is better than PCM, frequency modulation requires a narrower band-width than does pulse-code modulation. By comparison with PCM-AMt, frequency modulation for the optimal band requires, for $\rho > 22$, a larger band-width. It is interesting to compare the noise stability of FM and PCM-AMt for equal band-widths. Figure 2 shows the results of error computations in systems with FM for the frequency band equal to the optimal band for PCM-AMt, with $\alpha_p = 1$ and $\alpha_p = 2$. The curves show that, for identical frequency bands, FM is more stable under noise than is PCM-AMt ($\alpha_p = 1$) for $7 < \rho < 50$, and PCM-AMt ($\alpha_p = 2$) for $7 < \rho < 80$.

Thus, for each form of simple binary code, there is a certain range of values of ρ in which systems with frequency modulation are more stable under noise and require narrower band-widths.

We consider an example. Let it be given that $U_s = 50$ millivolts, $\sigma = \frac{2 \text{ millivolts}}{\sqrt{\text{cycle}}}$ and $T = 0.1$ second. It is required to determine the mean square error and the corresponding signal frequency bands for AM, FM and PCM. We find the magnitude of generalized parameter ρ from Formula (4). Since $F_m = 1/2T$, $\rho = 11.2$.

According to (5), the error for AM will equal $\delta_{AM} = 12.6\%$.

The frequency band for AM equals $\Delta f_{AM} = 2F_m = 10$ cycles. With this band-width, the signal-to-noise ratio at the receiver input will equal

$$(U_s / U_n)_{in\ AM} = U_s / \sigma \sqrt{\Delta f_{AM}} = 7.9.$$

To compute the noise stability with FM, we determine the optimal value of the parameter γ and the optimal deviation frequency. According to (33), $\gamma_{opt} = 7.8$.

We take $\gamma = 8$ and, correspondingly, $f_d = 40$ cycles. With this, the frequency band-width taken by the signal will be

$$\Delta f_{FM} = 2F_d + 2F_m = 90 \text{ cy.}$$

From (30) we determine the magnitude of the mean square error with FM:

$$\delta^2_{FM} = 20.2 \times 10^{-6}.$$

From whence,

$$\delta_{FM} = 0.45\%.$$

With a 90-cycle band-width, the signal-to-noise ratio at the FM-receiver input will be

$$\left(\frac{U_s}{U_n}\right)_{in\ FM} = \frac{U_s}{\sigma \sqrt{\Delta f_{FM}}} = 2.6.$$

We now determine the noise stability of PCM with secondary amplitude modulation (i.e., by the presence or absence of a code pulse) and with the intervals between pulses the same length as the pulses. With this, $\nu = 1$, $\alpha_p = 2$ and β, determined from Formula (21), equals 0.25.

The product $\beta\rho$, equals 2.8. Consequently, it is not possible to use Formulas (25) and (26) for determining the optimal number of pulses in the code, or the error. By computation using Formula (24), it is easily found that $n_{opt} = 3$. With this, the squared error equals $\delta^2_{PCM} = 0.0017 + 0.428V(1.62)$.

From tables,* we find that $V(1.62) = 0.053$. Hence, $\delta_{PCM\text{-}AMt} = 15.6\%$. The band-width required for

---

* Tables of the function $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ are given, for example, in the book of B. V. Gnedenko, "Course in Probability Theory." [In Russian].

PCM-AMt for $\mu = 1.3$ [see (27)] is $\Delta f_{PCM-AMt} = 156$ cycles. The signal-to-noise ratio at the PCM-AMt receiver will equal

$$(U_s / U_n)_{in\ PCM-AMt} = \frac{U_s}{\sigma \sqrt{\Delta f_{PCM-AMt}}} = 1.6.$$

We now determine the noise stability of PCM with secondary frequency keying and without intervals, i.e., for $\nu = \sqrt{2}$ and $\alpha_p = 1$.

In this case, $\beta = 0.5$, $\beta\rho = 5.6$ and, for the determination of n_{opt} and $\delta_{PCM-FMt}$, one can use Formulas (25) and (26): $n_{opt} \approx 4$ and $\delta_{PCM-FMt} = 4\%$. Band-width, $\Delta f_{PCM-FMt} = 208$ cycles. At the input of the PCM-FMt receiver, the signal-to-noise ratio is

$$(U_s / U_n)_{in\ PCM-FMt} = 1.7.$$

Thus, with the data considered, FM gives better noise stability than PCM-AMt and PCM-FMt, and requires a narrower band-width. In comparison with amplitude modulation, FM requires, for the optimal deviation, 9 times the band-width, but gives an error which is less by a factor of 29.

It sometimes happens in practice that, in FM systems with very low speeds of action (on the order of a second), it becomes difficult to attain the optimal band-width due to frequency instability of the transmitter and the frequency characteristics of the receiver, and deviation frequencies greater than optimal occur. In such a case, an FM system may turn out to have less noise stability than a PCM system, in which it is generally easier to attain an optimal band-width. However, in such slowly acting systems, the error due to noise is generally very small, with the signal-to-noise ratio at the receiver input greater than two, so that the requirements can also be satisfied by a system with $\gamma > \gamma_{opt}$. Thus, for example, in an FM system with $f_d = 75$ cycles, $T = 0.5$ seconds and $(U_s/U_n)_{in} = 2.3$, the error is 0.24%. With this, $\gamma = 75$. The optimal value of γ for the given system is 40, i.e., $f_{d\ opt} = 40$ cycles. Computations show that the error for $\gamma = \gamma_{opt}$ and the same values of signal, specific noise and speed of action, would equal 0.04%, i.e., would be six times smaller. If PCM-FMt ($\nu = \sqrt{2}$, $\alpha_p = 1$) were used for parameter transmission, then the mean square error for the same data would amount to 0.06%, i.e., would be somewhat larger than for FM with $f_d = 40$ cycles, but would be four times less than for FM with $f_d = 75$ cycles.

In this work, noise stability for PFM, PTM and PWM was considered only for weak noise. This hindered comparisons of them with other types of modulation. The results of analysis of noise stability of PTM, PWM and PFM with relatively strong noise fluctuations will be published in subsequent papers. In these will be given the limits of applicability of the formulas for weak noise for these forms of modulation, and comparisons will be made of PTM, PWM and PFM with FM and PCM.

LITERATURE CITED

[1] V. A. Kotel'nikov, Theory of Potential Noise Stability, (Gosenergoizdat, 1956) [In Russian].

[2] V. A. Kashirin, "Noise stability of frequency modulation systems," Automation and Remote Control (USSR) 18, 6 (1957).

[3] G. A. Shastova, "Investigation of noise stability of tele-control command transmission by the method of potential noise stability theory, parts I and II," Automation and Remote Control (USSR) 16, 4 (1955) and 17, 5 (1956).

[4] Iu. I. Chugin, "The optimal deviation frequency in single-channel telemetering systems," Automation and Remote Control (USSR) 19, 4 (1958).

Received October 8, 1957

LINEAR TRANSFORMATIONS OF BINARY CODES

N. Ia. Matiukhin

(Moscow)

Linear transformations of binary codes with fixed number of elements per code are considered such, that there is a one-to-one correspondence between any symbol and an object of coding (nondegenerate codes).*

Parallel to the binary number system with its wide field of applications in digital electronic computers, telemetry and telecontrol, a large number of other binary codes with their own distinct characteristics is also known. Here belong Hamming's self-correcting code [1], Gray code [2] for converters of continuous quantities into discontinuous ones, or the channel capacity increasing Fano-Shannon code [3], etc.

The methods of separately obtaining each of these single codes, as well as their properties, are described in detail in literature. However, much less attention has been paid to methods of construction applicable to a wider class of codes.

One of these methods based on representing coding symbols as vectors in a multi-dimensional space, and on investigations of linear transformations in that space is considered below.

We shall assume that numbers are the only objects of coding. To every number M from the set of numbers permissible in a given system there corresponds an encoding symbol of $\underline{m}$ elements $\alpha_1, \alpha_2, \ldots, \alpha_m$, (m-tuple), where each element of the symbol α_i can take only one of two values (0 or 1). The total number of different numbers which can be obtained using a binary m-tuple is equal to the number of different encoding symbols:

$$N_{\max} = C_m^0 + C_m' + \ldots, \quad C_m^{m-1} + C_m^m = (1+1)^m = 2^m, \tag{1}$$

where C_m^i is the number of selections of $\underline{i}$ elements out of $\underline{m}$.

The number of different correspondences (codes) between 2^m numbers and 2^m m-tuple code symbols is equal to the number of all possible rearrangements of symbols $N_m = 2^m!$. When m = 4, they already number 10^{13}, approximately.

We are choosing the 2^m numbers M to be the first integers, starting from zero: $0, 1, 2, \ldots, 2^m - 1$.

To any number M shall correspond m-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ by $M = \sum_{i=1}^{m} \alpha_i 2^{i-1}$, that is by making use of the binary representation of number M.

The code thus obtained is called normal code X_n.

---

* Self-correcting codes, for example, do not belong to this group, as in this case one object of coding corresponds to several encoding symbols.

Any other possible code X can be written in the form of a substitution table establishing a correspondence between the normal and the other code, for example when m = 2

$$\begin{pmatrix} 00, & 01, & 10, & 11 \\ 10, & 11, & 00, & 01 \end{pmatrix}, \quad \begin{pmatrix} 00, & 01, & 10, & 11 \\ 11, & 10, & 01, & 00 \end{pmatrix}, \quad \text{in general} \begin{pmatrix} X_n \\ X \end{pmatrix}.$$

Linear Substitutions Subgroup and Its Characteristics

We shall study some properties of the interesting substitution group. To do this we shall consider the m-dimensional vector space over field P_2, the residues modulo 2 [4]. In this space any vector is in the form $x = (x_1, x_2, \ldots, x_m)$, where x_i is equal either to 0 or 1. The addition of two vectors and the multiplication of a vector by a number from field P_2 is given by formulas

$$\mathbf{x} + \mathbf{y} = (x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_m \oplus y_m)\,^*,$$

$$\lambda \mathbf{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_m).$$

Null vector $\mathbf{0}$, such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$, is of the form $\mathbf{0} = (0, 0, \ldots, 0)$; vector $\mathbf{x}'$ opposite to $\mathbf{x}$, i.e., such that $\mathbf{x} + \mathbf{x}' = \mathbf{0}$ is equal to vector $\mathbf{x}$ itself $(\mathbf{x} + \mathbf{x} = \mathbf{0})$.

The following $\underline{m}$ linearly independent vectors are selected in our space:

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, 0, \ldots, 0, 0), \\ \mathbf{e}_2 &= (0, 1, 0, \ldots, 0, 0), \\ \mathbf{e}_3 &= (0, 0, 1, \ldots, 0, 0), \\ &\ldots\ldots\ldots\ldots\ldots \\ \mathbf{e}_m &= (0, 0, 0, \ldots, 0, 1), \end{aligned}$$

and the set of these vectors is called the base of our vector space. Any vector $\mathbf{x}$ can be obtained as a linear combination of the base vectors:

$$\mathbf{x} = (x_1, x_2, \ldots, x_m) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_m\mathbf{e}_m.$$

In the m-dimensional space there are 2^m different vectors and therefore a code symbol of a binary code may be made to correspond to each vector. We agree to write the symbol corresponding to a given vector in such a way that the i-th element of the symbol from the left corresponds at the same time to the i-th component of the vector and to the i-th coordinate in the base $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$.

For briefness we shall also denote a vector by its code symbol remembering it, however, when performing operations on vectors. For example, instead $\mathbf{x} = (1, 0, 0, 1)$ we shall write simply $\mathbf{x} = 1001$.

The vectors of our space can be subjected to linear transformation represented by square matrices of order $\underline{m}$, having all elements either zero or unity.

The operation of linear transformation is denoted in the form $\mathbf{y} = A\mathbf{x}$, where $\mathbf{x}$ is the input vector, and $\mathbf{y}$ the vector corresponding to it via transformation given by matrix A.

Transformations are performed in accordance with the usual matrix by vector multiplication rule. For example, if

$$A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}, \quad x = 101,$$

then $\mathbf{y} = A\mathbf{x} = 1 \cdot 1 \oplus 1 \cdot 0 \oplus 1 \cdot 1,\ 1 \cdot 1 \oplus 0 \cdot 0 \oplus 1 \cdot 1,\ 1 \cdot 1 \oplus 0 \cdot 0 \oplus 0 \cdot 1 = 001.$

* The sign $\oplus$ means addition mod 2.

The determinant of matrix A is found by applying usual rules with one modification viz. in any formula for the evaluation of the determinant all negative signs are changed to positive.*

Multiplication and addition of matrices are also performed according to known rules, and have properties analogous to those of common matrices.

If the determinant of matrix A is not zero (matrix A nondegenerate), there exists an inverse matrix A^{-1} such that the product $A \cdot A^{-1} = E$, where E is the unitary matrix. Later, we shall be mainly interested in nondegenerate matrices.

We shall consider the following two examples of nondegenerate transformations.

1. Identical transformation $E\mathbf{x} = \mathbf{x}$

$$E = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

2. The interchange of rows transformation

$$A = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$$

which interchanges the second with the fourth row. It is obvious that $A^{-1} = A$, i.e., $A^2 = E$.

How the nondegenerate transformations determine the one-to-one correspondence between the input vector $\mathbf{x}$ and the transformed vector $\mathbf{y}$ can be seen by regarding the vectors $\mathbf{x}$ and $\mathbf{y}$ as code symbols and using A to form a substitution table:

$$\begin{pmatrix} X_n \\ AX_n \end{pmatrix} = \begin{pmatrix} 000, & 001, & 010, & 011, \ldots \\ A(000), & A(001), & A(010), & A(011), \ldots \end{pmatrix}.$$

Substitutions of the type $\begin{pmatrix} X_n \\ AX_n \end{pmatrix}$, where A is any nondegenerate matrix, are fully determined by their matrix A. Such substitutions shall be called linear substitutions.

Any code obtainable by linear substitution shall be called linearly dependent on the normal code. The number of such codes (including the normal code) is equal to the number of different matrices A:

$$N_A^m = 2^{\frac{m(m-1)}{2}} \prod_{i=1}^{m} (2^i - 1).$$

The above formula is derived in the Appendix. The number of different matrices when m = 4 is already $N_A^4 \approx 2.2 \cdot 10^4$.

It has been shown above that all the possible 2^m substitutions form a finite group. One can show that all the linear substitutions form a subgroup of that group as all the possible nondegenerate matrices form a group closed with respect to matrix multiplication. The substitution group can therefore be decomposed into $N = 2^m!/N_A^m$ (right-hand or left-hand) cosets relative to that sub-group [5]. A coset is determined by one of its codes (in the substitution table). All other codes of that coset may be obtained by linear substitutions carried out on that code.

* Subtraction in P_2 field is equivalent to addition as $x = -x$.

Let us have, for example, a substitution $\begin{pmatrix} X_n \\ X \end{pmatrix}$, not belonging to the linear substitution sub-group. By means of this substitution a coset $\begin{pmatrix} X_n \\ AX \end{pmatrix}$ can be constructed. The linear sub-group also forms a coset generated by identical substitution $\begin{pmatrix} X_n \\ X_n \end{pmatrix}$ Therefore, any nondegenerate code may be formed by one of N codes (substitutions tables) defining N different cosets, and by one of N_A^m matrices from the linear sub-group. In this work, however, as mentioned above, our consideration is limited only to codes of the linear sub-group.

Later, it will be necessary to form matrix A from a linear substitution table. Matrix A is fully given by the simultaneous equations:

$$\begin{aligned} y_1 &= Ax_1, \\ y_2 &= Ax_2, \\ &\cdots\cdots \\ y_m &= Ax_m, \end{aligned}$$

where vectors $x_1, x_2, \ldots, x_m$ and also $y_1, y_2, \ldots, y_m$, entering the substitution table in twos, are linearly independent. Otherwise either matrix A would be degenerate or not determined, due to an insufficient number of equations.

If instead of vectors $x_1, x_2, \ldots, x_m$ we take normal base vectors $e_1, e_2, \ldots, e_m$, then vectors $y_1, y_2, \ldots, y_m$ are identical with columns of matrix A. If

$$A = \begin{vmatrix} a_{11}a_{12} \cdot \cdot a_{1m} \\ a_{21}a_{22} \cdots a_{2m} \\ \cdots\cdots\cdots \\ a_{m1}a_{m2} \cdots a_{mm} \end{vmatrix}$$

then

$$\begin{aligned} y_1 &= a_{11}a_{21} \cdots a_{m1}, \\ y_2 &= a_{12}a_{22} \cdots a_{m2}, \\ &\cdots\cdots\cdots\cdots \\ y_m &= a_{1m}a_{2m} \cdot\cdot a_{mm} \end{aligned} \tag{2}$$

Examples of Encoding Systems of the Linear Sub-group and Physical Realization of Linear Transformation Matrices

Example 1. We shall construct transformation A_1 such that the even symbol of code X_n and code Y be identical and that the odd X_n symbol have in Y a symbol obtained as a complement of its predecessor. The X_n symbols are listed according to Formula (1). When m = 4, our rules can be formulated in the following way:

$$A_1x_{2n} = y_{2n} = x_{2n}, \quad A_1x_{2n+1} = y_{2n+1} = x_{2n} \oplus 1111$$

or

$$A_1(x_{2n} \oplus x_{2n+1}) = 1111.$$

But $x_{2n} \oplus x_{2n+1} = 0001$ (they differ only in the least significant digit), that is

$$A_1(0001) = 1111.$$

One arranges the other three conditions for A_1 as follows:

$$A_1(0010) = 0010, \quad A_1(0100) = 0100, \quad A_1(1000) = 1000.$$

It is not difficult to verify that these four rules satisfy the requirement for vector linear independence, as stated above.

Using Relations (2), we obtain matrix

$$A_1 = \begin{vmatrix} 1001 \\ 0101 \\ 0011 \\ 0001 \end{vmatrix}$$

and the corresponding table

$$\begin{pmatrix} 0000, & 0001, & 0010, & 0011, & 0100, & 0101, \ldots, & 1110, & 1111 \\ 0000, & 1111, & 0010, & 1101, & 0100, & 1011, \ldots, & 1110, & 0001 \end{pmatrix}.$$

Example 2. We shall construct transformation A_2 such that to the adjacent symbols x_n and x_{n+1} of code X_n shall correspond y_n and y_{n+1} of code Y, differing from one another in one place only. First we shall find how symbols x_n and x_{n+1} differ from one another, in other words, what is their sum $x_n \oplus x_{n+1}$.

Symbol x_{n+1} is an arithmetical sum of x_n and the symbol which has 1 only in the least significant place. When adding, the carry-over 1 may have to be carried several places to the left and, depending on how many places it has been carried, the sum of interesting symbols may assume one of the following values:

$$x_n + x_{n+1} = \begin{cases} \ldots 0001, \\ \ldots 0011, \\ \ldots 0111, \\ \ldots 1111. \end{cases}$$

As assumed above,

$$y_n \oplus y_{n+1} = e_i,$$

where e_i is the symbol which has unity only in the i-th place, that is a normal base vector.

In view of $A_2(x_n \oplus x_{n+1}) = y_n \oplus y_{n+1}$, we can write down the system of equations for A_2 (when m = 4):

$$A_2(0001) = 0001, \quad A_2(0011) = 0010, \quad A_2(0111) = 0100, \quad A_2(1111) = 1000. \tag{3}$$

From these equations the matrix of the inverse transformation A_2^{-1} can be determined straightaway, writing these as follows:

$$0001 = A_2^{-1}(0001), \quad 0011 = A_2^{-1}(0010), \quad 0111 = A_2^{-1}(0100), \quad 1111 = A_2^{-1}(1000).$$

Hence, in agreement with (2), we have

$$A_2^{-1} = \begin{vmatrix} 1000 \\ 1100 \\ 1110 \\ 1111 \end{vmatrix}.$$

To find A_2 it is convenient to form another system of equations by adding the equations in (3) in two's in such a way that the base vector transforms appear on the left-hand side. In this way, we obtain

$$A_2(0001) = 0001, \quad A_2(0010) = 0011, \quad A_2(0100) = 0110, \quad A_2(1000) = 1100.$$

Hence,

$$A_2 = \begin{vmatrix} 1000 \\ 1100 \\ 0110 \\ 0011 \end{vmatrix}.$$

The corresponding substitution table takes the form

$$\begin{pmatrix} 0000, & 0001, & 0010, & 0011, & 0100, & 0101, & 0110, & 0111, & \ldots, & 1111 \\ 0000, & 0001, & 0011, & 0010, & 0110, & 0111, & 0101, & 0100, & \ldots, & 1000 \end{pmatrix}.$$

The code thus obtained is called Gray code and is used with a quantizer transforming continuous quantities into discontinuous ones by using a code wheel or a plate. Figure 1 shows a schematic outline of a shift-digit converter with a Gray code wheel. The angular position of the shaft is converted into a number with an accuracy of 1/16th of a revolution. The dotted places on the wheel are electric conductors. The number corresponding to the position of the wheel is brought down from contact 1-4 in the form of a binary digit. Voltage is brought to the wheel through a central contact. On each contact 1-4 voltage corresponds to digit 1, no voltage to digit 0.

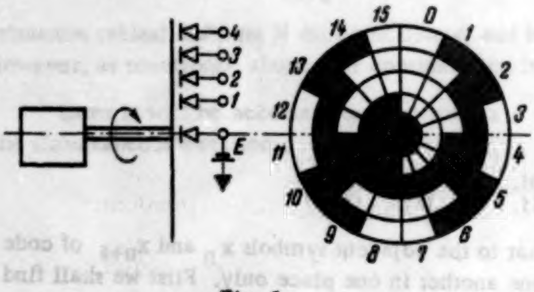


Fig. 1

While the voltage in the contacts is changing the error in the determination of the wheel position, which may occur in the normal code, is eliminated because the adjacent positions on the wheel differ in one place only. For instance, normal code symbols corresponding to numbers 7 and 8 differ in all their places and during transition from position 7 to position 8 any symbol corresponding to one of the positions 1-16 might be obtained, depending on the diameters of the contacts 1-4 and how well they were fixed.

Figure 2 shows a photograph of a wheel with 512 positions, obtained photographically.

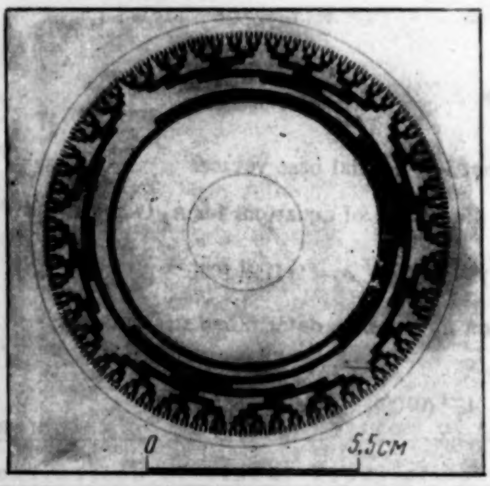


Fig. 2

It should be noted that though it follows from (3) that there exist m! different transformations A_2, which can be obtained by rearranging the right-hand side of (3) in different ways, the codes thus obtained differ from Gray code only by transposition of places and therefore are of little interest to us.

The inverse transformation, from Gray code into the normal code performed by matrix A_2^{-1}, can have a very simple technical realization.

The converter (Fig. 3) has trigger T fed through its counter successively – place after place – by Gray code symbols in the form of series of pulses, starting from the top digit. The sum of digits (modulo 2) of the higher places of Gray code symbol appears consecutively on the trigger. The sum of the $\underline{k}$ highest places of Gray code symbol gives the k-th place of the normal code symbol if our adding starts with the top digit. The normal code symbol is being obtained place after place from trigger T through valve K. It is easy to verify that this code transformation corresponds to matrix A_2^{-1}.

The demerit of the converter is that the symbols are obtained starting from the top digits on the left of the binary representation, and that these symbols cannot therefore be used directly in automatic digital computers where the feeding of the code symbol always starts with the least significant digit on the right of the binary representation.

In the example below properties will be considered of a code whose normal code symbols are obtained by using a device shown in Fig. 3, which begins with the least significant digit of the binary system.

Example 3. We shall examine the properties of a code whose inverse transform matrix is given in the form

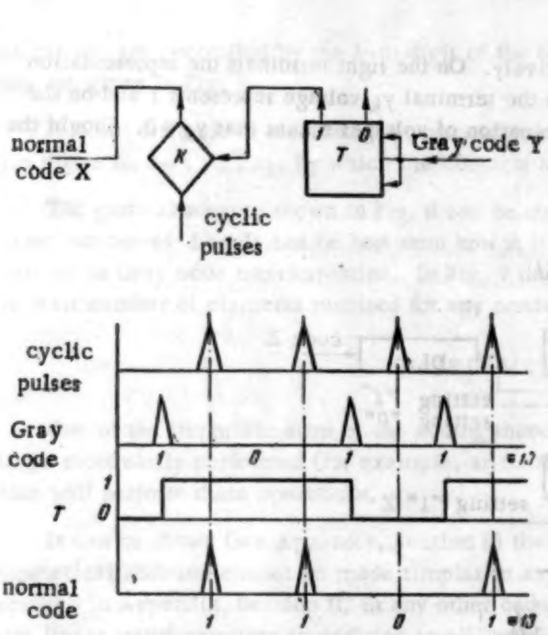


Fig. 3

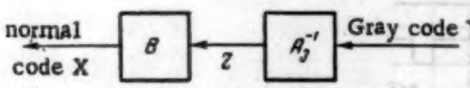


Fig. 4

$$A_3^{-1} = \begin{vmatrix} 1&1&1&1 \\ 0&1&1&1 \\ 0&0&1&1 \\ 0&0&0&1 \end{vmatrix}.$$

But matrix A_3^{-1} is the transpose of A_2^{-1} matrix in our last example, hence matrix A_3 is obtained as a transpose of A_2 matrix of the same example:

$$A_3 = \begin{vmatrix} 1&1&0&0 \\ 0&1&1&0 \\ 0&0&1&1 \\ 0&0&0&1 \end{vmatrix}$$

Transformation A_3 leads to relations:

$$A_3(0001) = 0011,\ A_3(0011) = 0101,$$
$$A_3(0111) = 1001,\ A_3(1111) = 0001. \qquad (4)$$

In this way any two consecutive y_n and y_{n+1} are distinguished by the least significant digit and differ in not more than one place. Comparing (4) with (3) we find that the symbols of system Y have – with the exception of the least significant place – the same structure as Gray code symbols.

Example 4. We shall describe a method of converting Gray code symbols into normal by using, as an intermediate device, the converter shown in Fig. 3 which is fed with Gray code symbols starting with the least significant digit. The converter in this case performs the transformation corresponding to the matrix in Example 3:

$$A_3^{-1} = \begin{vmatrix} 1111 \\ 0111 \\ 0011 \\ 0001 \end{vmatrix}$$

We shall find an additional transformation B such that when applied after A_3^{-1} (Fig. 4) the normal code will result.

As $z = A_3^{-1}y$ and $x = Bz$, then $x = BA_3^{-1}y = A_2^{-1}y$.

Hence $BA_3^{-1} = A_2^{-1}$ and $B = A_2^{-1}A_3$, that is when $m = 4$

$$B = \begin{vmatrix} 1&0&0&0 \\ 1&1&0&0 \\ 1&1&1&0 \\ 1&1&1&1 \end{vmatrix} \cdot \begin{vmatrix} 1&1&0&0 \\ 0&1&1&0 \\ 0&0&1&1 \\ 0&0&0&1 \end{vmatrix} = \begin{vmatrix} 1&1&0&0 \\ 1&0&1&0 \\ 1&0&0&1 \\ 1&0&0&0 \end{vmatrix}. \qquad (5)$$

It can be seen from (5) that transformation B is such that top digit z_1, being at the same time the least significant digit x_4, forms the digit of the required symbols z_4, z_3, z_2 when added successively to z_1 modulo 2.

The realization of this transformation is shown in Fig. 5.

We shall now consider relay contact networks suitable for obtaining a realization of any transformation matrix (Fig. 6). In the case of active zero action each matrix row represents one switch circuit which gets voltage

on its left input terminals representing digits 1 and 0, respectively. On the right terminals the representation of the digit y_k of the transformed symbol y is obtained; if on the terminal y_k voltage represents 1 and on the terminal $\bar{y}_k$ it represents 0 then $y_k = 1$. The reciprocal combination of voltages means that $y_k = 0$. Should the

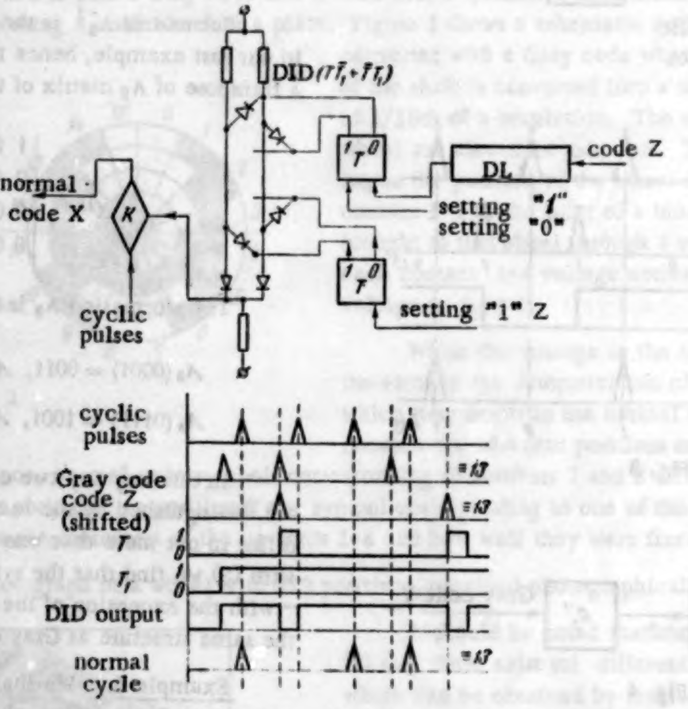


Fig. 5. DID — disaccord decoder, DL — delay line for (n + 1) places.

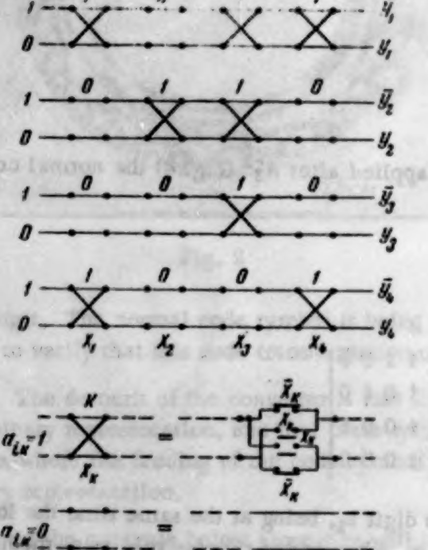


Fig. 6. $A = \begin{vmatrix} 1011 \\ 0110 \\ 0010 \\ 1001 \end{vmatrix}$

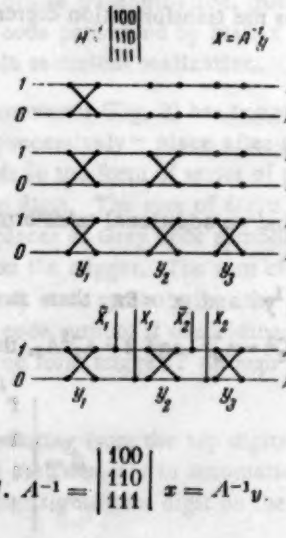


Fig. 7. $A^{-1} = \begin{vmatrix} 100 \\ 110 \\ 111 \end{vmatrix}$ $x = A^{-1}y$.

0 digit be represented by the breaking of the circuit, the lower left and the upper right terminals can be left out. The i-th row switch circuit is expressed by successive couplings of the contact circuit corresponding to the unitary elements in this matrix row. The contacts of any

such circuit are controlled by the k-th digit of the code being transformed. The network diagram and the usual terms are given in Fig. 5.

It is not difficult to ascertain that successive couplings of such circuits form a network which adds modulo 2 to the digits $x_1, x_2, \ldots, x_k$, by which the contacts are controlled.

The general scheme shown in Fig. 6 can be simplified considerably when the columns of matrix A contain a large number of 1's. It can be best seen how it is done in the example of the construction of the matrix which is inverse to Gray code transformation. In Fig. 7 the general and simplified schemes of this converter are given. The least number of elements required for any nondegenerate matrix cannot be less than its order $\underline{m}$.

SUPPLEMENTARY REMARKS

One of the important aims of the coding theory is the construction of codes on which various operations will be most easily performed (for example, arithmetic operations, etc.) as well as the construction of algerithms which will perform these operations.

It can be shown (see Appendix, Section I) that the operation which in the normal code corresponds to an arithmetical addition cannot be made simpler in any other code of the linear sub-group of binary codes. In fact, as proved in Appendix, Section II, in any other code (of the linear sub-group) it is necessary to perform three extra linear transformations in addition to all the elementary operations necessary for adding in the normal code.

As far as the self-correcting codes are concerned, an example can be cited of linear transformations which lead from the normal to the self-correcting code. For instance, the transformation

$$A = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{vmatrix}$$

converts a four-digit normal code symbol into a seven-digit Hamming code symbol which corrects a single error. The inverse problem, however, the determination of the normal code symbol cannot be solved by linear transformation methods as the above matrix has no reciprocal. The solution of this problem requires more involved methods.

At the same time it is possible, for example, to construct a single error-detecting code. Let $\mathbf{x}$ be the symbol of such a code and $\mathbf{y} = A\mathbf{x}$. We shall only use those $\mathbf{x}$ symbols which have $x_1 = 0$. Requiring an error in one single place to produce one in x_1 place of the $\mathbf{x}$ symbol, we obtain (when m = 4)

$$A^{-1}(0001) = 1001, \quad A^{-1}(0010) = 1010,$$

$$A^{-1}(0100) = 1100, \quad A^{-1}(1000) = 1000$$

$$\text{or} \quad A^{-1} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad ; \text{ hence} \quad A = A^{-1} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

Looking at the matrix it is easy to see that the code obtained is the one which is known to check a single error by the sum of digits. In a general case, it is possible to construct a code with required characteristics by the method of code linear transformations, only when one is able to present these characteristics in the form of a one-to-one correspondence between linearly independent vectors of a given code system on the one hand and the required code system on the other.

It is obvious from the above that the number of such coding systems is large and increases quickly with m.

APPENDIX

1. We shall determine the number of different nondegenerate transformation matrices, by applying mathematical induction.

Let N_A^{m-1} different transformations be in the $(m-1)$ dimensional space. We shall compute the number of matrices A_m of order m of the following type:

$$A_m^* = \left|\begin{array}{c|c} A_{m-1} & \begin{matrix}0\\0\\:\\0\\0\\0\end{matrix} \\ \hline B & 1 \end{array}\right|$$

Here A_{m-1} is a nondegenerate matrix, order $(m-1)$.

For any matrices A_{m-1} and B the matrix A_m^* is nondegenerate, because the columns of A_{m-1} are linearly independent and so are, therefore, the first $m-1$ columns of matrix A. Also the m-th column must be linearly independent of the others because there is no linear combination of columns of A_{m-1} which could give zero elements in the m-th column (this would be against the condition of linear independence). It follows also that matrices A_m^* formed by all matrices A_{m-1} and all B matrices exhaust all possible nondegenerate transform matrices A_m which have the above m-th column.

It is clear that the number of such matrices is

$$N = N_A^{m-1} N_B,$$

where N_B is the number of different B matrices. But B can be represented in the binary system by a number with $(m-1)$ digits and hence $N_B = 2^{m-1}$.

Thus, the number of A_m^* matrices with specified m-th column has been found. As we are free to fill in the m-th column in any way, with just the one exception of all zeros, and as with each such column goes the same number of A_m^* matrices as obtained originally, the total number of different transformations A_m is therefore

$$N_A^m = (2^m - 1)\,N = (2^m - 1)\,2^{m-1} N_A^{m-1}.$$

As $N_A^1 = 1$ when $m = 1$, we obtain

$$N_A^m = \prod_{i=1}^{m} (2^i - 1) \times 2^{i-1} = 2^{\frac{m(m-1)}{2}} \prod_{i=1}^{m} (2^i - 1).$$

II. Performing arithmetical addition with transformed symbols.

Let the arithmetic addition operation $x_1 + x_2 = x_3$ be performed on normal code symbols x_1 and x_2 and let these be such that x_3 has the same dimensionality.

It is required to find such an operation on the mapped symbols y that $y_1 0 y_2 = y_3$, where $y_1 = Ax_1$, $y_2 = Ax_2$, $y_3 = Ax_3$.

We know that

$$x_1 + x_2 = x_1 \oplus x_2 \oplus u, \tag{6}$$

where u is the binary carry-over code symbol. Hence

$$y_3 = y_1 \oplus y_2 \oplus Au, \tag{7}$$

which means that to be able to describe the arithmetical addition with transformed codes it is necessary to determine the transformed carry-over symbol Au. We shall show that Au cannot be expressed by y_1 and y_2 unless the inverse mapping A^{-1} is used.

To do this we shall represent **u** as $u = u_a \oplus u_b$, where u_a is the symbol of unconditional carry-over and u_b the symbol of conditional carry-over.

To compute u_a and u_b it is necessary to introduce, apart from addition modulo 2, yet another operation on the symbols; this operation is the place-by-place or logical multiplication $x = x_1x_2$. This operation is not linear with respect to transformation A, that is

$$A(x_1x_2) \neq Ax_1Ax_2. \tag{8}$$

The unconditionsl carry-over symbol is obviously $u_a = M(x_1x_2)$, where

$$M = \begin{vmatrix} 010 \ldots 00 \\ 001 \ldots 00 \\ \cdots \\ 000 \ldots 10 \\ 000 \ldots 01 \\ 000 \ldots 00 \end{vmatrix}$$

is the shift mapping transferring the obtained carry-over into the next place.

The conditional carry-over vector is

$$u_b = u_1 \oplus u_2 \oplus u_3 \ldots,$$

where $u_k = M[u_{k-1}(x_1 \oplus x_2)]$ and $u_1 = M[u_a(x_1 \oplus x_2)]$.

The first term of the sum u_1 corresponds to the conditional carry-over resulting from the unconditional one occurring in the preceding place. The second and following terms represent conditional carry-overs, each obtained from its predecessor, as seen from the last two formulas.

Thus

$$u = M[(x_1x_2) \oplus M[(x_1x_2)(x_1 \oplus x_2)] \oplus M\{[M[(x_1x_2)(x_1 \oplus x_2)]](x_1 \oplus x_2)\} \ldots]. \tag{9}$$

Hence, it follows from Construction (9) and Inequality (8) that to express the transformed symbol Au (7) in terms of y_1 and y_2, it is necessary to calculate first vector **u** itself which can be done when y_1 and y_2 are known by substituting $x_1 = A^{-1}y_1$ and $x_2 = A^{-1}y_2$ into Formula (9). As seen above,if the result of arithmetical addition of code symbols y_1 and y_2 is to be obtained in the Y code — following remarks on computing vector **u** and Au and using Formulas (6) and (7) — the same intermediate operations must be performed: computation of vector **u**, two inverse transformations A^{-1}, one A transformation and two additions modulo 2.

Further, if the result of the addition — expressed in the normal code — can be used then it is not necessary to perform transformation A when using Formula (6). Comparing arithmetical addition operation in the normal code (6) and any other code (7) it can be seen that in the last case it is necessary to perform three additional linear transformations: $A^{-1}y_1$, $A^{-1}y_2$, and Au.

LITERATURE CITED

[1] R. W. Hamming, Error detecting and error correcting codes, The Bell System Technical Journal 26, 2 (1950).

[2] N. J. Gray, P. V. Levonian and M. Rubinoff, An analog-to-digital converter for serial computing machines, Proc. IRE 10 (1953).

[3] C. E. Shannon, Statistical theory of transmission of electrical signals [Russian translation].

[4] B. L. Van der Waerden, Modern Algebra, Vol. 1 (1957), [Russian translation].

[5] A. G. Kurosh, Course of Higher Algebra, (Nat. Techn. Press, 1955) [In Russian].

Received September 17, 1957.

STATIC CHARACTERISTICS OF TOROIDAL CORES AS A FUNCTION OF THEIR GEOMETRIC DIMENSIONS

M. A. Rozenblat

(Moscow)

The paper treats of the influence of the ratio of the inside and outside diameters of toroidal cores on their magnetic characteristics and, in particular, on the fundamental magnetization curves and the forms of the hysteresis loops. Formulas are given for computing these characteristics. It is shown that core geometry is most essential in cases where magnetic materials are employed which have rectangular hysteresis loops or sharp corners on the fundamental magnetization curve.

INTRODUCTION

Toroidal cores have found extensive employment in magnetic amplifiers, saturated chokes, dc measuring transformers, magnetic relays, magnetic memory devices, and in many devices where it is necessary to make maximum use of the magnetic properties inherent in the material of which the cores are made. The advantages of toroidal cores are that they allow practical exclusion of the deleterious influence of air gaps and leakage, and permit a more uniform magnetization to be obtained than do other types of cores. However, even toroidal cores do not allow complete uniformity of magnetization over the entire core to be obtained. This is a result of the differing magnetic path lengths along the inside and outside perimeters, thanks to which the intensity of the field generated by the current in the windings varies across the core's cross section.

The lack of uniformity of magnetization along the toroidal core's cross section frequently leads to a significant deterioration in its magnetic properties, compared to the properties of the material forming the core.

The present paper is devoted to a consideration of the influence of the ratio of outside and inside diameters of toroidal cores on their magnetization curves and the form of their hysteresis loops. We remark that this question was partially treated in work [1]. The authors of the work cited limited themselves to explaining the influence of the core dimensions on the steepness of the descending arm of the limiting hysteresis loop in cases where alloys with rectangular hysteresis loops were used. With this, only the central (linear) parts of the descending arm of the hysteresis loop were considered. Here, this question is considered more fully.

Magnetic Characteristics of Cores

In what follows, we shall denote by $\underline{b}$ and $\underline{h}$ the magnetic induction and field strength at a given point of a core, and by B and H, the average values of magnetic induction and field strength of the core. Corresponding with this, the relationship, $b = f(h)$, is the magnetic characteristic (magnetization curve or hysteresis loop) of the material, and the functional relationship, $B = F(H)$, is the magnetic characteristic of the core.

The fundamental assumption made in this work is that the core is taken from a homogeneous body, for any point of which the relationship, $b = f(h)$, is valid.

We denote by R_O and R_I, respectively, the outer and inner radii of the core, and by α their ratio, $\alpha = R_O/R_I$. Then, the average induction of the core is

$$B = \frac{1}{R_I - R_I} \int_{R_I}^{R_O} b\, dR, \tag{1}$$

where $\underline{b}$ is the magnitude of the induction at the point situated at a distance R from the geometric center of the core.

Analogously, we find the average field strength of the core:

$$H = \frac{1}{R_O - R_I} \int_{R_I}^{R_O} h\, dR.$$

If the current I, in the winding $\underline{w}$, placed around the core, generates a magnetizing force $F = 0.4\,\pi\,wI$, then

$$h = \frac{F}{2\pi R}. \tag{2}$$

By substituting this value of $\underline{h}$ in the equation for H, we find that

$$H = \frac{F \ln R_O/R_I}{2\pi(R_O - R_I)} = \frac{F \ln \alpha}{2\pi R_I(\alpha - 1)}. \tag{3}$$

Thus, the average core field strength equals the field strength at points lying on a circle with radius

$$R_h = \frac{R_O - R_I}{\ln R_O/R_I} = \frac{R_I(\alpha - 1)}{\ln \alpha}, \tag{4}$$

called the mean harmonic radius.

We mention that, in practice, the core field strength is frequently computed by dividing the magnetizing force by the average core length:

$$H' = \frac{F}{l_{av}} = \frac{F}{\pi(R_O + R_I)}, \tag{5}$$

which gives a lower value for the average field strength than does Formula (3).

We provide several values of the ratio H/H', computed from the formula

$$\frac{H}{H'} = \frac{\alpha + 1}{2(\alpha - 1)} \ln \alpha.$$

α	1.4	1.7	2	3	4	5
H/H'	1.01	1.02	1.04	1.11	1.16	1.21

For values of $\alpha \leq 2$, it is possible for engineering computations to calculate the average core field strength by the simpler Formula (5) without any essential error.*

In the future in calculating H we will be using Formula (3).

* In work [1], the authors used Formula (5) even for higher values of α which led to implausible results. For example, it turned out that an increase of α led in some cases to an improvement in the core's magnetic characteristics, which does not, in fact, occur.

Core Magnetization Curves

The core magnetization curve can be computed if one has an analytic expression for the material magnetization curve $b = f(h)$, or can be found graphically if one has the function $b = f(h)$ in the form of an experimental curve.

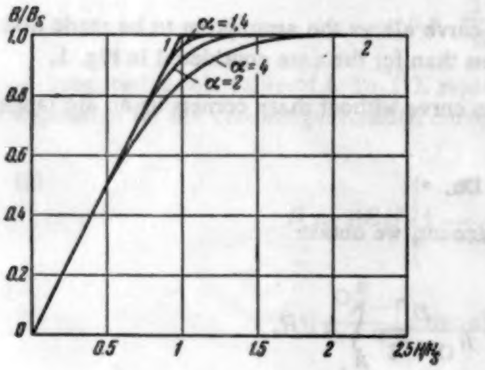


Fig. 1

We first consider the case when the material magnetization curve is represented by the broken-line curve 0-1-2 (Fig. 1). We denote by B_s the saturation induction and by H_s the field strength for which the core saturates, giving us

$$b = \begin{cases} B_s \dfrac{h}{H_s} & \text{for } h \leqslant H_s \\ B_s & \text{for } h \geqslant H_s. \end{cases} \tag{6}$$

If there are no saturated portions in the core, i.e., if $h < H_s$, then by substituting the value of $\underline{b}$ from (6) in Equation (1), we find the average core induction:

$$B = \frac{B_s}{H_s}\,\frac{1}{R_O - R_I}\int_{R_I}^{R_O} h\,dR = \frac{B_s}{H_s}\,H.$$

For $\dfrac{F}{2\pi R_I} = H_s$, which, according to Formula (3), corresponds to an average value of the core field strength of

$$H = H_k = H_s\,\frac{\ln\alpha}{\alpha - 1}, \tag{7}$$

the inside layer of the core, having radius R_O, is saturated. Therefore, for $H > H_k$, it is necessary to compute the core induction by the formula

$$B = \frac{1}{R_O - R_I}\left[(R_p - R_I)\,B_s + \int_{R_p}^{R_O} \frac{B_s}{H_s}\,h\,dR\right],$$

where R_p is the radius of the circle dividing the core into saturated and unsaturated sections:

$$R_p = \frac{F}{2\pi H_s} = \frac{R_I(\alpha - 1)}{\ln\alpha}\,\frac{H}{H_s}.$$

If we substitute Expression (2) for $\underline{h}$, and take (3) into account, we find, after integrating, the following expression for B when $H \geq H_k$:

$$B = \frac{B_s}{\alpha - 1}\left[\frac{H}{H_s}\,\frac{\alpha - 1}{\ln\alpha}\left(1 + \ln\frac{\alpha\ln\alpha}{\alpha - 1}\,\frac{H_s}{H}\right) - 1\right] \tag{8}$$

Thus, for $H \leq H_k$, the core magnetization curve coincides with the material magnetization curve, and B varies in proportion to H. For $H > H_k$, the core magnetization curve becomes nonlinear, and may differ significantly from the material magnetization curve.

We present the values of H_k, computed from Formula (7), for which the linearity of the core magnetization curve is disrupted.

α	1	1.4	1.6	1.8	2	3	4
H_k/H_s	1	0.84	0.78	0.73	0.69	0.55	0.46

Figure 1 gives, in relative units, the core magnetization curves computed by Formula (8) for $\alpha = 1.4$, 2, and 4. If, for $\alpha = 1.4$, the influence of the inhomogeneity of core magnetization is small, for $\alpha = 2$ and $\alpha = 4$ this influence becomes significant, and engenders a noticeable shortening of the linear portion of the magnetization curve and a smoothing out of the sharp corners of the curve. We note that the relative influence of the magnitude of α on the magnetization curve does not depend on the magnitude of the magnetic permeability (i.e., on the ratio B_s/H_s).

The character of the influence of α on the magnetization curve allows the assumption to be made that for materials without sharp corners, the influence of α will be less than for the case considered in Fig. 1.

It is frequently convenient to approximate a magnetization curve without sharp corners by an arc tangent curve, sometimes with a linear term:

$$b = A \ \mathrm{arc\ tg}\ Ch + Dh. \ * \tag{9}$$

By substituting this expression in (1), and taking (2) into account, we obtain

$$B = \frac{A}{R_O - R_I} \int_{R_I}^{R_O} \mathrm{arctg}\, \frac{CF}{2\pi R}\, dR + \frac{D}{R_O - R_I} \int_{R_I}^{R_O} h\, dR.$$

The second term in the right member equals DH. To simplify the first integral, we introduce a new variable:

$$x = CF/2\pi R.$$

Then, the integral takes the following form [3]:

$$\int \mathrm{arctg}\, \frac{CF}{2\pi R}\, dR = -\frac{CF}{2\pi} \int \frac{\mathrm{arctg}\, x}{x^2}\, dx = -\frac{CF}{2\pi} \left[\ln x - \frac{1}{2} \ln (1 + x^2) - \frac{\mathrm{arctg}\, x}{x}\right].$$

Replacing F by H in accordance with Formula (3), we find, after setting in the limits of integration:

$$B = (AC + D) H + \frac{ACH}{2 \ln \alpha} \ln \frac{1 + \left(CH \frac{\alpha - 1}{\alpha \ln \alpha}\right)^2}{1 + \left(CH \frac{\alpha - 1}{\ln \alpha}\right)^2} + A\left(\frac{\alpha}{\alpha - 1}\, \mathrm{arc\ tg}\, CH \frac{\alpha - 1}{\alpha \ln \alpha} - \frac{1}{\alpha - 1}\, \mathrm{arc\ tg}\, CH \frac{\alpha - 1}{\ln \alpha}\right). \tag{10}$$

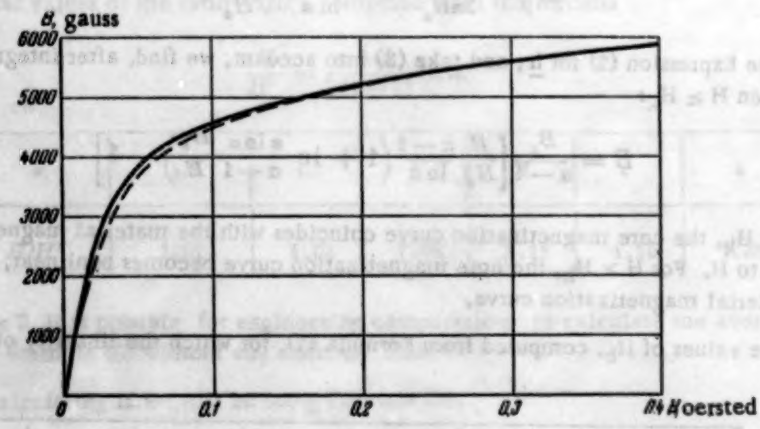


Fig. 2

As an example, we consider the magnetization curve $b = f(h)$, shown on Fig. 2 and computed by Formula (9) with the values A = 3190 gauss, C = 45/oersted and D = 2400 gauss/oersted. This curve, for values of $h \le 0.5$ oersted, corresponds with high accuracy to the magnetization curve of alloy 80NKhS with maximum permeability

* 'tg' = 'tan', 'sh' = 'sinh', — Publisher.

μ_{max} = 138,000 gauss/oersted and coercive force H_c = 0.019 oersted (see [2], p. 541). The dotted curve is for the core magnetization curve B = F(H), computed by Formula (10) with $\alpha = 4$. Even for such a large value of α there is a relatively small worsening of the magnetization curve.

In certain cases, the material magnetization curves are approximated by hyperbolic sines:

$$h = \frac{1}{C}\ \mathrm{sh}\ \frac{b}{A} \quad \text{or} \quad b = A\ \mathrm{arc\ sh}\ Ch. \tag{11}$$

Substituting this value of h in (1), replacing R by the new x = 1/R, and integrating, we find an expression for the core magnetization curve in the form:

$$B = ACH\left\{1 - \frac{1}{\ln\alpha}\ln\frac{1+\sqrt{1+\left[\frac{CH(\alpha-1)}{\ln\alpha}\right]^2}}{1+\sqrt{1+\left[\frac{CH(\alpha-1)}{\alpha\ln\alpha}\right]^2}}\right\} + \frac{A}{\alpha-1}\left[\alpha\ \mathrm{arc\,sh}\,\frac{CH(\alpha-1)}{\alpha\ln\alpha} - \mathrm{arc\,sh}\,\frac{CH(\alpha-1)}{\ln\alpha}\right]. \tag{12}$$

Computation of the core and material magnetization curves B = F(H) and b = f(h), respectively, by Formulas (11) and (12) also showed a relatively small difference between them for $\alpha = 4$.

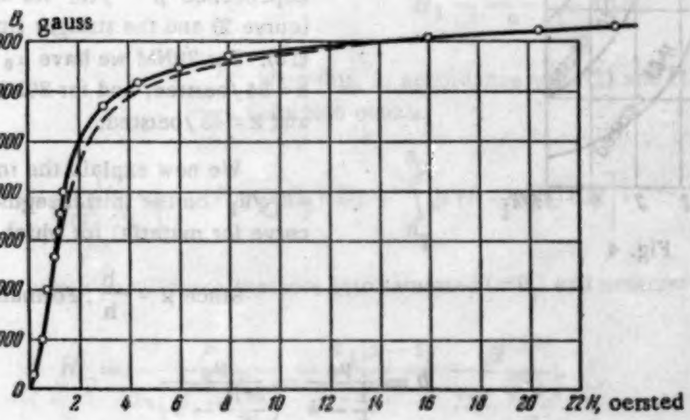


Fig. 3

If the relationship b = f(h) is in the form of an experimental curve, the core magnetization curve can be found by graphically integrating Equation (1). For this, the values of H should be given, and the magnitude of the field strength h should be calculated at several points of the core situated at various distances R from its center. For these h, the corresponding values of b are found from the curve b = f(h), and the relationship b = φ(R) is constructed for the given H. The area included between the axis of abscissas, the vertical lines R = = R_I and R = R_O, and the curve b = φ(R) equals (when the scales are taken into account) the integral in the right member of (1). By repeating such a construction for various values of H, we find the desired functional relationship B = F(H).

As an example, Fig. 3 shows, by the solid line, the measured fundamental magnetization curve b = f(h), for electrotechnic steel E42, and Fig. 4 shows the curves b = $\varphi(R/R_I)$, computed by the method described above for various values of H for a toroid with $\alpha = 5$. After graphical integration, the core magnetization curve B = = F(H), shown by the dotted line on Fig. 3, was obtained.

Formulas (9) and (11) do not always assure a sufficiently accurate approximation of the fundamental magnetization curve in the regions of weak fields. For many magneto-soft materials, the magnetic permeability in these regions increases in direct proportion to the magnitude of the magnetic induction, and may be represented as

$$\mu = \mu_0 + kb, \tag{13}$$

where μ_0 is the initial permeability.

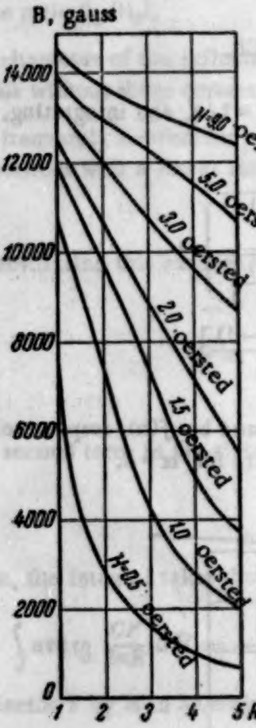


Fig. 4

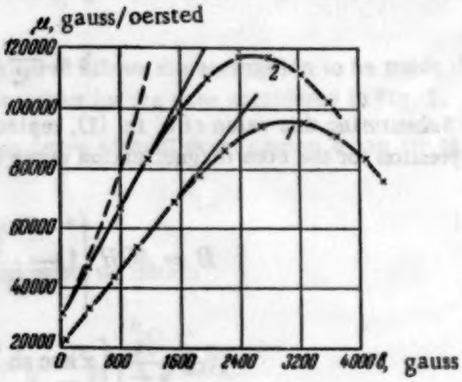


Fig. 5

As an illustration, Fig. 5 shows the experimental dependence $\mu = f(b)$ for alloy 79NM (curve 1), 80NKhS (curve 2) and the straight lines corresponding to Formula (13). For 79NM we have $\mu_0 = 20{,}500$ gauss/oersted and $k = 34$/oersted, and for 80NKhS, $\mu_0 = 30{,}000$ gauss/oersted and $k = 45$/oersted.

We now explain the influence of the quantity $\alpha = R_O/R_I$ on the initial segment of the core magnetization curve for material for which Formula (13) is valid.

Since $\mu = \frac{b}{h}$, Formula (13) can be put in the form:

$$b = \frac{\mu_0}{\frac{1}{h} - k} = \frac{\mu_0}{\frac{2\pi R}{F} - k}$$

and, for the mean core induction, we find that

$$B = \frac{1}{R_O - R_I}\int_{R_I}^{R_O} \frac{\mu_0 \frac{F}{2\pi}}{R - \frac{kF}{2\pi}}\, dR = \frac{\mu_0 F}{2\pi (R_O - R_I)} \ln \frac{R_O - \frac{kF}{2\pi}}{R_I - \frac{kF}{2\pi}}.$$

By expressing F in terms of the field strength H in accordance with (3), we find that

$$B = \frac{\mu_0 H}{\ln \alpha} \ln \frac{\alpha \ln \alpha - kH(\alpha - 1)}{\ln \alpha - kH(\alpha - 1)}. \tag{14}$$

The dotted line on Fig. 5 shows the initial portion of the curve $\mu = B/H = F(B)$, computed from Formula (14) for a core made of alloy 80NKhS ($\mu_0 = 30{,}000$ gauss/oersted and $k = 45$/oersted) with $\alpha = 4$. It is clear that for $\alpha > 1$ there is an increased core permeability in weak fields, as compared with the permeability of the material.

If we denote by h_k the maximum value of field strength for which Formula (13) remains valid then, obviously, we can use Formula (14) for values

$$H \leqslant h_k \frac{\ln \alpha}{\alpha - 1}.$$

Core Hysteresis Loops

The hysteresis loop of a toroidal core, account being taken of the nonuniformity of magnetization along the cross section, is also computed by means of Formula (1).

We shall first consider materials with a so-called "rectangular" hysteresis loop, represented in somewhat idealized form in Fig. 6. The residual induction B_r is taken to equal the saturation induction B_s. The ascending arm of the hysteresis loop is described by the equation:

$$b = \mu (h - H_c), \tag{15}$$

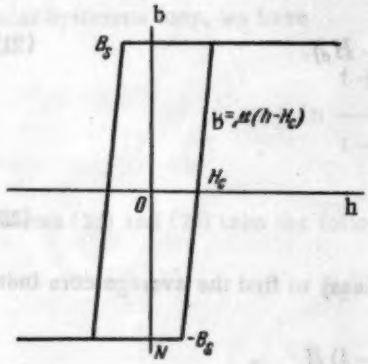


Fig. 6

where

$$H_c - B_s/\mu \leqslant h \leqslant H_c + B_s/\mu.$$

We shall assume that the core's initial magnetic state corresponds to point N of Fig. 6. Therefore, for $h \le H_C - B_s/\mu$ we have $b = -B_s$ and, for $h \ge H_C + B_s/\mu$, the induction is $b = B_s$. The core begins to reverse its magnetic polarity, from state $-B_s$ to state B_s, when the core's average field strength H reaches the value:

$$H_1 = \frac{H_c - \frac{B_s}{\mu}}{\alpha - 1} \ln \alpha. \tag{16}$$

With this, in accordance with (1) and (15), the mean value of core induction equals:

$$B = \frac{1}{R_O - R_I}\left[-(R_O - R_p) B_s + \int_{R_I}^{R_p} \mu (h - H_c)\, dR\right],$$

where R_p is the radius of the circle partitioning the core into saturated $(-B_s)$ and unsaturated parts:

$$R_p = \frac{F}{2\pi\left(H_c - \frac{B_s}{\mu}\right)} = \frac{R_I(\alpha - 1)}{\ln \alpha} \frac{H}{H_c - \frac{B_s}{\mu}}.$$

After integration, we get the following expression for B, for $H \ge H_1$:

$$B_1 = \frac{\mu H_c - \alpha B_s}{\alpha - 1} + \frac{\mu H}{\ln \alpha}\left[\ln\left(\frac{\alpha - 1}{\ln \alpha} \frac{H}{H_c - \frac{B_s}{\mu}}\right) - 1\right]. \tag{17}$$

For $H \ge H_2$, where

$$H_2 = \frac{H_c + \frac{B_s}{\mu}}{\alpha - 1} \ln \alpha, \tag{18}$$

the portion of the core adjoining the inner perimeter $(R = R_I)$ has completely reversed its magnetic polarity, and has an induction of $b = B_s$.

For $H = H_3$, where

$$H_3 = H_1 \frac{R_O}{R_I} = \left(H_c - \frac{B_s}{\mu}\right) \frac{\alpha \ln \alpha}{\alpha - 1}, \tag{19}$$

all portions of the core have begun to reverse polarity, i.e., for $H > H_3$, at any point of the core $b > -B_s$. Therefore, Formula (17) is only valid for values of $H \le H_2$ and $H \le H_3$. The further course of the functional

relationship $B = F(H)$ is determined by the relationship between H_2 and H_3. Two cases are possible.

1. $H_3 < H_2$. This case corresponds to

$$\alpha < \frac{1 + \frac{B_s}{\mu H_c}}{1 - \frac{B_s}{\mu H_c}}. \tag{20}$$

If $H_3 \le H \le H_2$, there are no saturated portions of the core, and

$$B_2 = \frac{\mu}{R_O - R_I} \int_{R_I}^{R_O} (h - H_c)\, dR = \mu (H - H_c). \tag{21}$$

For $H_2 \le H \le H_4$, where

$$H_4 = H_2 \frac{R_O}{R_I} = \left(H_c + \frac{B_s}{\mu}\right) \frac{\alpha \ln \alpha}{\alpha - 1}, \tag{22}$$

is the field strength corresponding to complete core polarity reversal, it is easy to find the average core induction:

$$B_3 = -\frac{B_s + \mu H_c \alpha}{\alpha - 1} + \frac{\mu H}{\ln \alpha} \left[1 - \ln \frac{(\alpha - 1) H}{\alpha \ln \alpha \left(H_c + \frac{B_s}{\mu}\right)}\right]. \tag{23}$$

Thus, if Condition (20) holds, the ascending arm of the hysteresis loop consists of a linear (central) portion (21) and two nonlinear portions, (17) and (23), at the beginning and end of the arm.

2. $H_3 > H_2$. In this case

$$\alpha > \frac{1 + \frac{B_s}{\mu H_c}}{1 - \frac{B_s}{\mu H_c}}. \tag{24}$$

For $H_2 \le H \le H_3$, the core has a portion with positive saturation (B_s), adjacent to the inner perimeter, a portion with negative saturation ($-B_s$), adjacent to the outer perimeter and an intermediate, nonsaturated portion. Therefore, the average core induction is determined from the equation:

$$B_2 = \frac{1}{R_O - R_I} \left[B_s (R_1 - R_I) - B_s (R_O - R_2) + \int_{R_1}^{R_2} \mu (h - H_c)\, dR \right],$$

where

$$R_1 = \frac{(\alpha - 1) R_I H}{\left(H_c + \frac{B_s}{\mu}\right) \ln \alpha} \text{ and } R_2 = \frac{(\alpha - 1) R_I H}{\left(H_c - \frac{B_s}{\mu}\right) \ln \alpha}.$$

By integrating, and then substituting the values for R_1 and R_2, we obtain

$$B_2 = -B_s \frac{\alpha + 1}{\alpha - 1} + \frac{\mu H}{\ln \alpha} \ln \frac{1 + \frac{B_s}{\mu H_c}}{1 - \frac{B_s}{\mu H_c}}. \tag{25}$$

For $H_3 \le H \le H_4$, the core induction B_3 is determined from Formula (23).

Thus, if Condition (24) holds, the central portion of the ascending (and the descending) arm of the hysteresis loop is characterized by a linear dependence of the induction on the field strength (25). However, unlike the

previous case (20), the differential core permeability on this portion,

$$\mu_{core} = \frac{dB}{dH} = \frac{\mu}{\ln \alpha} \ln \frac{1 + \frac{B_s}{\mu H_c}}{1 - \frac{B_s}{\mu H_c}} \qquad (26)$$

does not equal the differential material permeability, and may be significantly less.

It is interesting to note that, for $\mu H_C \gg B_s$ and, in particular, for $\mu \to \infty$, which corresponds to the ideal rectangular hysteresis loop, we have

$$\ln \frac{1 + \frac{B_s}{\mu H_c}}{1 - \frac{B_s}{\mu H_c}} = \ln \left(1 + 2\frac{B_s}{\mu H_c}\right) = 2\frac{B_s}{\mu H_c}$$

and Equations (25) and (26) take the following forms:

$$B_2 = B_s\left(-\frac{\alpha+1}{\alpha-1} + \frac{2}{\ln \alpha}\frac{H}{H_c}\right), \qquad (25')$$

$$\mu_{core} = \frac{2}{\ln \alpha}\frac{B_s}{H_c}. \qquad (26')$$

With this, $H_2 = H_1$ and $H_3 = H_4$, i.e., the ascending and descending arms of the core's hysteresis loop consist of just one linear portion (25').

If we set $B_2 = 0$ in Formulas (25) and (25'), and compute the corresponding values of H, we obtain the magnitude of the core's coercive force $H_{C_{core}}$, which turns out to be larger than the material's coercive force.

For example, from Formula (25'), we have

$$H_{C_{core}} = \frac{(\alpha+1)\ln \alpha}{2(\alpha-1)} H_c.$$

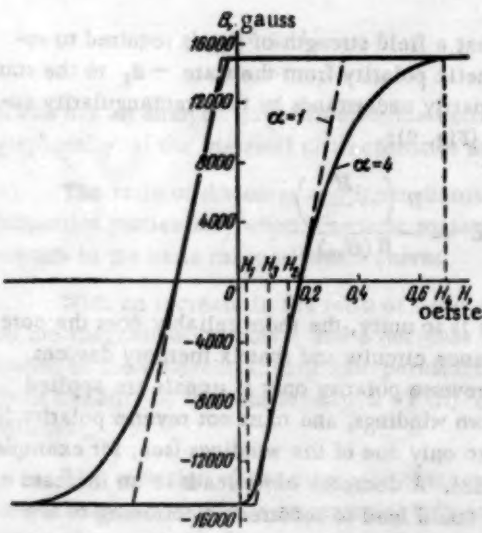


Fig. 7

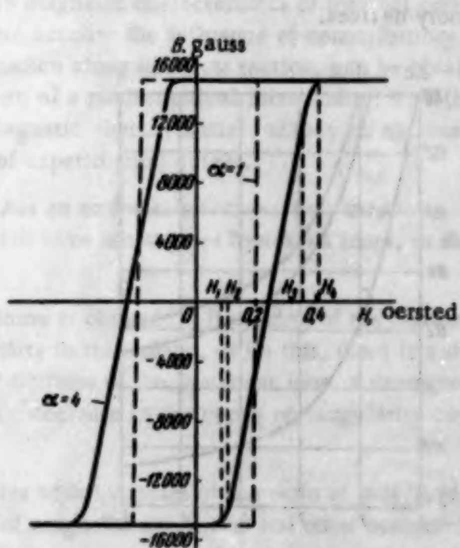


Fig. 8

If the average core field strength is computed by Formula (5), then its coercive force turns out to be:

$$H'_{C_{core}} = H_C.$$

Figures 7 and 8 give the static hysteresis loops for toroidal cores with $\alpha = 4$, computed by the formulas given above. The dotted lines represent the material hysteresis loops. The hysteresis loop of Fig. 7 corresponds to the first case considered above (B_s = 15,000 gauss, H_c = 0.2 oersted and $\mu = 10^5$ gauss/oersted), and the loop of Fig. 8 corresponds to the second case (B_s = 15,000 gauss, H_c = 0.2 oersted and $\mu = 10^6$ gauss/oersted).

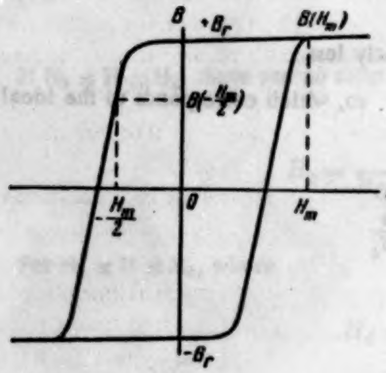


Fig. 9

It is obvious from Fig. 7 that, if Condition (20) holds, the relationship $B = F(H)$, on the ascending and descending arms of the hysteresis loop for a core with a sizeable value of $\alpha = R_O/R_I$, differs significantly from a linear one. For example, for $\alpha = 4$, the induction region, where the relationship $B = F(H)$ is linear, consists of only 21% of the entire range of variation of the induction ($2B_s$).

The deviation of the relationship $B = F(H)$ from the linear engenders the same type of infringement of "input-output" characteristic linearity as certain types of magnetic amplifiers with positive feed-back [1, 2].

It is clear from Fig. 8 that, if Condition (24) holds, an increase in α shows its effects principally in a decrease in the differential permeability dB/dH on the "vertical" portions of the hysteresis loop of the core, which leads to a decrease in the gain and quality in the corresponding types of magnetic amplifiers with positive feed-back.

By substituting in Formula (26) the values, typical for alloy 50NP, of B_s = 15,000 gauss, H_c = 0.2 oersted and $\mu = 10^6$ gauss/oersted, we obtain the following values of μ_{core}:

α	1–1.16	1.4	2	4
μ_{core}, gauss/oerst.	10^6	4.5×10^5	2.16×10^5	1.1×10^5

The formulas given above can also be used to compute the rectangularity coefficient of a hysteresis loop, which is one of the important parameters characterizing the quality of cores used in magneto-static triggers and memory devices.

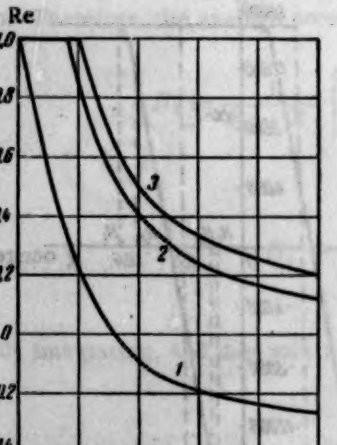


Fig. 10

We assume that a field strength of H_m is required to reverse a core's magnetic polarity from the state $-B_r$ to the state B_r. Then, one ordinarily understands by the "rectangularity coefficient" the ratio (Fig. 9):

$$\mathrm{Re} = \frac{B\left(-\frac{H_m}{2}\right)}{B(H_m)}.$$

The closer Re is to unity, the more reliably does the core function in coincidence circuits and matrix memory devices, where a core must reverse polarity only if signals are applied simultaneously to two windings, and must not reverse polarity if a signal is applied to only one of the windings (see, for example, [2], paragraphs 13-18). A decrease of Re leads to an increase of the noise level and could lead to incorrect functioning of the devices mentioned.

As applied to our idealized hysteresis loop (Figs. 6-8), $H_m = H_4$, $B(H_m) = B_s$ and

$$\mathrm{Re} = \frac{B\left(-\frac{H_4}{2}\right)}{B_s}. \tag{27}$$

Figure 10 shows the dependence of Re on α, computed by Formulas (17) - (25) and (27), for three values of $\mu H_C/B_s$. The curve ($\mu H_C/B_s = 3$) corresponds to an alloy of type 50NP and molybdenum Perminvar, with the values $H_C = 0.1$ oersted, $B_s = 15{,}000$ gauss and $\mu = 450{,}000$ gauss/oersted. Even for $\alpha = 1.5$, the rectangularity coefficient falls to 0.56.

An increase of H_C and a decrease of B_s (for example, by the use of ferrite cores) allows Re to be increased for a given value of α. For $\mu H_C/B_s = 15$ (curve 2), the rectangularity coefficient is close to the limiting value of Re, corresponding to $\mu = \infty$, and any further increase of the ratio, $\mu H_C/B_s$, leads only to a small increase of Re, for a given α. For $\mu = \infty$, we have Re = 1 if $\alpha \le 2$, and $\mathrm{Re} = \dfrac{1}{\alpha - 1}$ for $\alpha > 2$ (curve 3).

It should be expected, from a consideration of Figs. 1, 2 and 3, that the influence of the magnitude of α on the form of the limiting hysteresis loop of materials with a small ratio of residual induction B_r, to saturation induction B_s, will be significantly less than for materials with rectangular hysteresis loops.

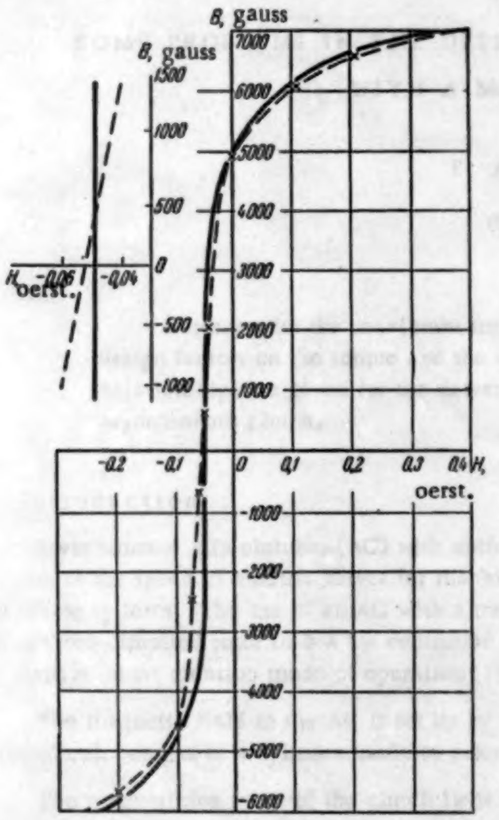


Fig. 11

On Fig. 11, the solid curve is the descending arm of a limiting hysteresis loop, obtained experimentally for a core of alloy 79NMA, with a value of $\alpha = 1.1$. The dotted curve was obtained by a graphical integration of Equation (1) for the value $\alpha = 4$.

It is of interest to note that if, in Formula (26), the value of B_r is substituted for that of B_s, this formula gives a value, $\mu_{core} = 146{,}000$ gauss/oersted (for $B_r = 5000$ gauss, $H_C = 0.05$ oersted and $\mu = 10^6$ gauss/oersted), which practically coincides with the value of μ_{core} obtained above.

SUMMARY

The magnetic characteristics of toroidal cores, taking into account the influence of nonuniformity of magnetization along the cross section, can be obtained in the form of a mathematical relationship, B = F (H), if one has an analytic expression for the corresponding material magnetic characteristics, or may be obtained graphically, if the material characteristics are given in the form of experimental curves.

The ratio of the outer and inner diameters of toroidal cores has an essential effect on their magnetic properties particularly when magnetic materials are employed which have rectangular hysteresis loops, or sharp corners in the basic magnetization curves.

With an increase in the ratio of outer and inner diameters, there is observed a flattening of the corners on the magnetization curve, and a decrease of magnetic permeability in this region. With this, there is a decrease of the differential magnetic permeability on the "vertical" portions of the hysteresis loop, a departure from linearity of the relationship B = F (H) on these portions, and a decrease in the core's rectangularity coefficient.

The variation of the magnetic characteristics of toroidal cores with increases of the ratio of outer and inner diameters can have an essential influence on the properties of magnetic amplifiers and other contact-less magnetic elements of automatic equipment, and must be taken into account in choosing the geometric dimensions of toroidal cores.

LITERATURE CITED

[1] R. W. Roberts and R. I. Van Nice, "Influence of the ID/OD ratio on the static and dynamic properties of toroidal cores," Trans. AIEE, 74, 1 (1955).

[2] M. A. Rozenblat, "Magnetic amplifiers," Sovetskoe Radio [In Russian] (1956).
[3] G. B. Dwight, Tables of Integrals [In Russian translation] (IL, 1948).

Received July 17, 1957

SOME PROBLEMS IN THE DESIGN OF AN ASYNCHRONOUS CLUTCH WITH A MONOLITHIC ROTOR

T. A. Glazenko

(Leningrad)

Formulas for the maximum torque and critical slip are deduced. The effect of various design factors on the torque and the shape of the mechanical characteristic is determined. Relationships are given for the determination of the optimum number of pole-pairs for the asynchronous clutch.

1. Introduction

Asynchronous slip clutches (AC) with uniform rotors find application in the automatic control and regulation of the speed of electric drives for machine tools and transport machines, and also as actuating mechanisms in tracking systems. The use of an AC with a monolithic rotor permits the utilization of asynchronous motors with a short-circuited rotor in 3-4 kw controlled installations with heavy starting conditions, shock loading, and a repetitive short duration mode of operation [1].

The magnetic field in the AC is set up by the dc winding situated on the inductor (Fig. 1). The armature of the clutch consists of a cylinder made of a ferromagnetic material.

The magnetizing force of the clutch Iw is constant and is independent of the relative rotor speed. Consequently, in setting up the electromagnetic field equations, the equivalent linear loading of the inductor Δ is taken as a basic quantity, and the field in the armature and gap of the AC is examined.

Equations for the clutch's maximum torque M_{cr}, which are obtained from the solution of the system of equations for the electromagnetic field in the gap and rotor of the clutch, are given in this paper.

The derived relationships make it possible to evaluate the effect on the torque and the form of the mechanical characteristic of the AC, of various design factors: the dimensions of the magnetic circuit — the outside diameter of the inductor D, the active length of pole l and air-gap δ cm, the relative magnetic permeability μ h/cm, conductivity of the rotor steel γ ohm^{-1} cm^{-1}, and the number of pole-pairs in the inductor p.

The inductor can be nonphaneropolar, phaneropolar, toothed with intersecting poles, and finger-toothed (Fig. 1).

An attempt is made in the paper to compare the operating properties of the AC with various types of inductors.

2. Clutch With a Nonphaneropolar Inductor

a. Linear loading of the inductor.

An inductor with not clearly defined poles makes it possible to obtain a close approximation to a sine wave in the curve of the distribution of the linear load Δ, and the normal component of induction B_θ, in the gap, for ideal lost motion.

The ampere-turns of one pair of poles of such an inductor, expressed in amps, are found from the formula

$$Iw_{\mathrm{p}} = 2\int_{0}^{\tau/2} \Delta_{\max}\cos\frac{\pi}{\tau}\cdot x\,dx = \frac{2\Delta_{\max}}{\pi} = \Delta_{\max}\frac{D}{p},$$

where $\underline{x}$ is the running coordinate on the evolute of the inductor's circumference.

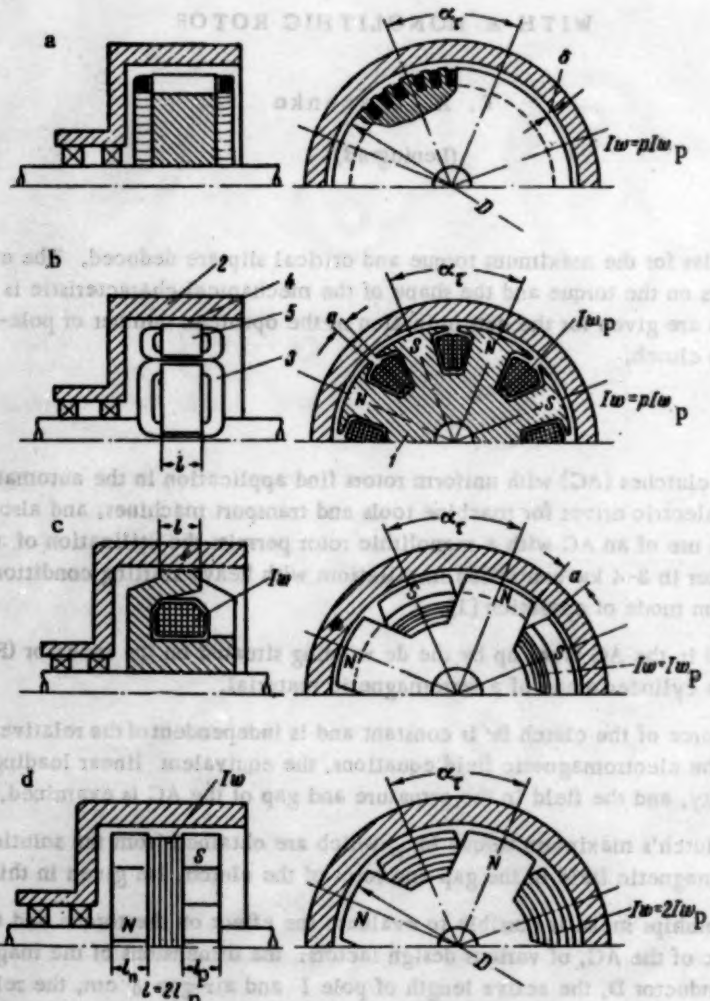


Fig. 1. Mechanical schematics of asynchronous clutches. a) Clutch with a nonphaneropolar inductor; b) with a phaneropolar inductor: 1) inductor, 2) rotor, 3) field winding, 4) gap, 5) pole; c) with a toothed inductor and alternating poles; d) with a toothed inductor.

Taking into account that the pole separation $\tau = \pi D/2p$, we obtain

$$Iw_{\mathrm{p}} = \Delta_{\max} D/p.$$

If we assume that the magnetic circuit of the clutch is unsaturated, then the amplitude of the nonlinear inductor load, expressed in amp/cm, can be expressed in the form

$$\Delta_{max} = \frac{Iw_p p}{D} = \frac{2B_{0max}\,\delta_e}{\mu_0}\,\frac{p}{D}. \quad (1)$$

Here $B_{0\,max}$ is the amplitude of the inductor in the gap, for ideal idle motion, when the angular frequency of the armature current $\omega_{el} = 0$, δ_e is the size of the equivalent air gap in the clutch, determined during the design of the magnetic circuit, $\mu_0 = 4\pi\ 10^{-9}$ h/cm is the magnetic permeability of air.

In the design of a clutch, the amplitude of the induction in the gap, for ideal idle motion, is first picked, $B_{0\,max}$ = 8000 - 12,000 gauss, and the necessary ampere-turns Iw_p are determined.

b. Torque and optimum number of pole-pairs.

The equations for the electromagnetic field in the gap and rotor of an AC with not clearly defined poles are set up under the usual assumptions [2, 3].

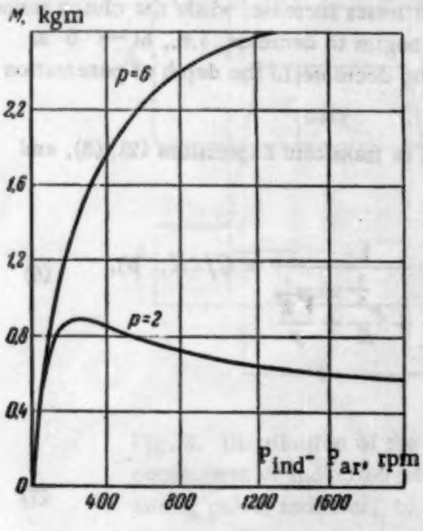


Fig. 2. Calculated mechanical characteristics for a nonphaneropolar AC with D = 13.4 cm, l = 3.0 cm, δ = 0.03 cm, Iw_p = 1500 amp, δ_e = 0.034 cm, $\mu = 10^3$, $\gamma_{80°C}$ = 0.312 × × 10^5 ohm^{-1} cm^{-1}.

The joint solution of Maxwell's equations yields the following expression for the torque on the armature due to the field:

$$M = \frac{(2+\sqrt{2})\,M_{cr}}{\sqrt{\frac{\omega}{\omega_{cr}}} + \sqrt{\frac{\omega_{cr}}{\omega}} + \sqrt{2}}, \quad (2)$$

where ω is the relative angular speed of the rotor:

$$\omega = \omega_{ind} s\ 1/\mathrm{sec}^{-1} \qquad s = \frac{\omega_{ind} - \omega_{ar}}{\omega_{ind}}$$

The relative angular speed for which the clutch torque reaches a maximum is:

$$\omega_{cr} = \omega_{ind} s_{cr} = \frac{16 p^3 \delta_e^2}{D^4 \gamma}\,\frac{\mu}{\mu_0}. \quad (3)$$

Here ω_{cr} and ω_{ar} are the angular speeds of rotation of the inductor and armature, s_{cr} is the critical slip.

The maximum (critical) torque is

$$M_{cr} = 0.266\pi\mu_0 \frac{D^3 l}{8p} \Delta^2_{max} A_\rho, \quad (4)$$

where A_ρ is a coefficient which takes into account the variation of the magnetic permeability in the direction of propagation of the electromagnetic wave in the rotor [4]; $A_\rho \approx 1.3 - 1.4$.

Formula (2) can be used for a qualitative evaluation and comparison of the shapes of the mechanical characteristics of different AC.

A more accurate plot of the curve $M = f(s)$ for an AC with a toothed inductor, within the range of operating speeds, can be obtained if one makes use of the formula given in [1].

Equation (2) indicates the presence of a maximum in the clutch's mechanical characteristic, which corresponds to physical phenomena and is confirmed experimentally (Figs. 2 and 4). In fact, for synchronous rotation of the inductor and armature (the mode for ideal idle motion $\omega_{el} = \omega_{ind}\,sp = 0$) eddy currents are not induced in the armature and the torque is equal to zero. However, if $\omega_{el} \to \infty$ then the depth of penetration of the wave into the mass of the rotor $z = \frac{1}{\sqrt{\gamma\mu\,\omega_{ind}\,sp}}$ vanishes together with the rotor losses and the torque.

The value of the relative speed $\omega_{ind} s_{cr}$ at which the maximum torque is reached depends on the dimensions of the clutch components (D, δ) and the electromagnetic characteristics (μ, γ) of the armature material. The value of ω_{cr} increases sharply with increase in the number of pole-pairs or the length of the air gap, as

well as with decrease of the clutch diameter. An increase in ω_{cr} with M_{cr} = const corresponds to a smoothing of the AC mechanical characteristic.

It should be noted that in clutches with a sharply nonsinusoidal distribution of the equivalent, linear load on the inductor in the gap, or with an appreciable number of pole-pairs, the maximum torque is usually reached for amounts of slip lying well outside the limits of the operating range of changes in the armature speed.

The expression for the critical torque follows from Relationships (1) and (4):

$$M_{cr} = 10.5 \times 10^{-9} \frac{Dl}{\delta} p (Iw_p)^2 A_p. \tag{5}$$

If the dimensions, ampere-turns Iw_p, and relative speed $\omega_{ind} s_n$ for which the nominal torque is reached, are chosen, then the most profitable number of pole-pairs p_{opt} can be determined.

An increase in the number of pole-pairs is accompanied by an increase in the electric speed of eddy currents in the rotor $\omega_{el} = \omega_{ind} sp$. At the same time the armature heat losses increase, while the clutch torque increases up to a definite limit. With further increase in p the torque begins to decrease, i.e., $M \to 0$ as $p \to \infty$, due to the ever growing effect of the armature reaction and the decrease in the depth of penetration of the electromagnetic wave into the rotor.

To determine the optimum number of pole-pairs it is convenient to transform Expressions (2), (3), and (5) to the form

$$M = 3.58 \times 10^{-8} \frac{DlA_p}{\delta} (Iw_p)^2 \frac{p}{\frac{K}{p^{\frac{3}{2}}} + \frac{p^{\frac{3}{2}}}{K} + \sqrt{2}} = C \frac{1}{\frac{K}{p^{\frac{5}{2}}} + \frac{p^{\frac{1}{2}}}{K} + \frac{\sqrt{2}}{p}} = Cf(K, p), \tag{6}$$

where

$$K = \frac{D^2}{4\delta} \sqrt{\frac{\gamma\mu_0}{\mu} \omega_{ind} s}. \tag{7}$$

The torque reaches a maximum when $\frac{\partial M}{\partial p} = 0$, or $2.5 K p^{-\frac{7}{2}} - \frac{0.5}{K} p^{-\frac{1}{2}} + \sqrt{2} p = 0$.

In this way, the optimum number of pole-pairs depends on the value of K which, in its turn, is determined by the clutch parameters.

The relative speed $\omega_{ind} s$ in (7) is chosen from the conditions for a stable transmission of torque. Usually one takes $\omega_{ind} \approx (1.5 - 3)\, \omega_{ind} s_n$.

The curve $p_{opt} = f(k)$ (Fig. 6, curve $a/D \to 0$) is used to determine the optimum number of pole-pairs for an inductor with not clearly defined poles.

3. Clutch With Phaneropolar and Toothed Inductor

a. Linear load on inductor.

For the case when $\omega_{el} = 0$, the distribution of induction in the gap, for a phaneropolar, toothed inductor with intersecting poles, has a trapezoidal form. The corresponding curve of the distribution of the equivalent linear load $\Delta = f(x)$ along the circumference of the inductor is given in Fig. 3 a.

The maximum value of the linear load on the inductor is $\Delta'_{max} = 2B_{0\,max}\, \delta_e / \mu_0 a$, where a is the calculated distance between the poles of the inductor. For a clutch with a phaneropolar inductor

$$\Delta'_{max} = \frac{Iw_p}{a}, \quad Iw_{tot} = p/w_p \tag{8}$$

For a clutch with a toothed inductor (with intersecting poles)

$$\Delta'_{max} = \frac{Iw_p}{a}, \quad Iw_{tot} = Iw_{p'} \tag{9}$$

where Iw_{tot} is the total ampere-turns packed on the inductor, a is the distance between the intersecting teeth.

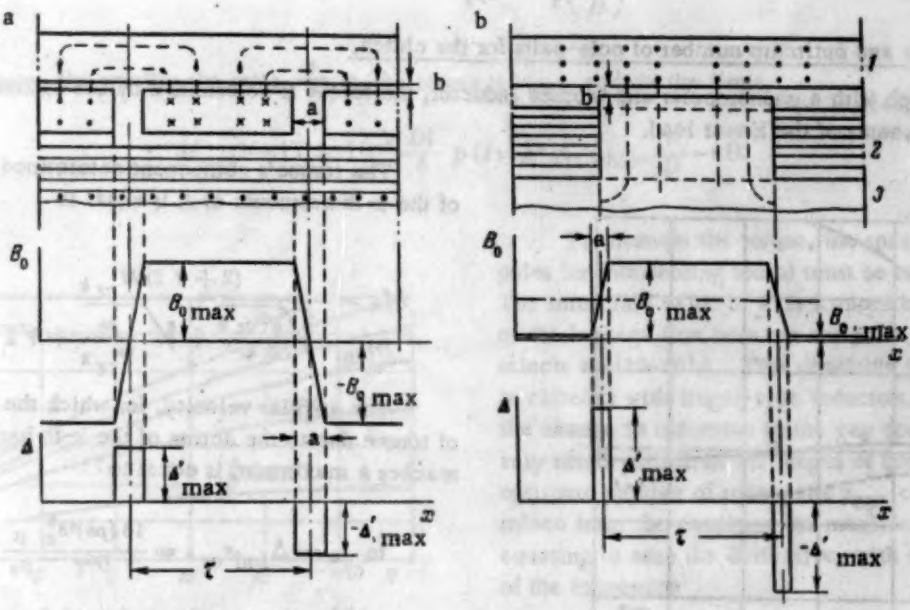


Fig. 3. Distribution of the equivalent linear load on the inductor and the normal component of induction when $\omega = 0$: a) for a phaneropolar and toothed (with alternating poles) inductor, b) for a toothed (finger-type) inductor. 1) Armature; 2) coil; 3) inductor.

To obtain a certain value of induction for idle motion, the clutch with toothed inductor must contain p times less ampere-turns than a clutch with clearly defined poles.

For a toothed (finger-type) inductor, the induction changes from $B_{0\,max}$ to $B_{0\,min}$ over the length of a very small portion of the gap's circumference (Fig. 3 b). Here $a = 0.1 - 0.3$ cm, usually.

In this case the maximum value of the equivalent linear load can be found from the equation

$$\Delta'_{max} = \frac{(B_{0\,max} - B_{0\,min})\,\delta_e}{\mu_0 a} \approx \frac{B_{0\,max}\,\delta_e}{\mu_0 a} = \frac{Iw_p}{a}, \tag{10}$$

where $Iw_p = \dfrac{Iw_{tot}}{2}$.

The active length of axis in the finger-type of inductor is $l = 2l_p$ (Fig. 1).

Ia. B. Rozman [1] proposed a diagonal displacement of poles in the finger-type of inductor. Such an arrangement weakens the edge effect on the inner sides of the poles, which results in an increased torque.

The distribution of the equivalent, linear load along the circumference of a toothed and phaneropolar inductor can be represented by a trigonometric series of the form

$$\Delta = \sum \Delta_{k\,max} \cos \frac{\pi}{\tau_k} x,$$

where $k = 1, 3, 5, \ldots$ – is the order number of the harmonic of Δ, $\Delta_{k\,max}$ is the amplitude of the k-th harmonic component of the curve $\Delta = f(x)$:

$$\Delta_{k\,max} = \frac{4\Delta'_{max}}{\pi k}\sin\frac{\pi}{\tau_k}\frac{a}{2} = \frac{4}{k\pi}Iw_p\frac{\sin\left(\frac{ap_k}{D}\right)}{a},$$

$$\tau_k = \frac{\tau}{k} = \frac{\pi D}{2p_k}, \quad p_k = kp. \tag{11}$$

b. Torque and optimum number of pole-pairs for the clutch.

For a clutch with a phaneropolar and toothed inductor, the torque is determined by the action of all the harmonic components of the linear load.

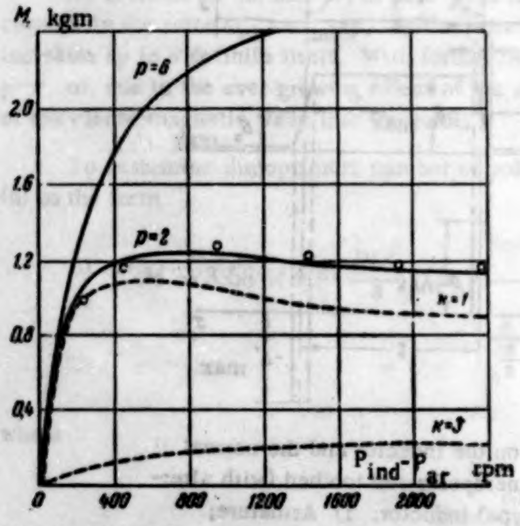


Fig. 4. Calculated and experimental characteristics of an AC with a phaneropolar inductor with D = 13.4 cm, δ = 0.04 sec, l = 3.0 cm, $\gamma_{80°C} = 0.312\cdot10^{6}$ ohm^{-1} cm^{-1}, $\mu = 10^{3}$, δ_e = 0.071 cm, Iw_p = 1500 amp. The circles denote experimental points.

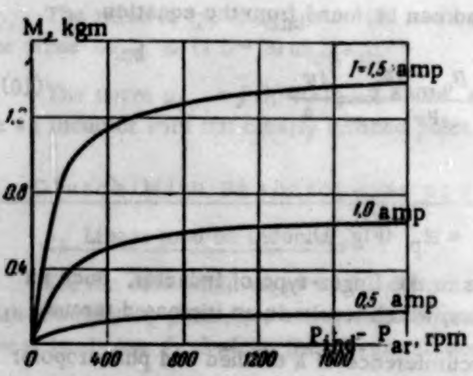


Fig. 5. Calculated mechanical characteristics of an AC with a toothed (finger-type) inductor, for D = 14.8 cm, $l = 2l_{p_1}$ = 5.6 cm, δ = 0.03 cm, $\gamma_{80°C} = 0.312\cdot10^{6}$ ohm^{-1} cm^{-1}, $\mu = 10^{3}$, δ_e = 0.034 cm, $p = p_{opt} = 6$, w_{tot} = 800 turns.

The torque's component determined by the action of the k-th harmonic of Δ is equal to

$$M_k = \frac{(2+\sqrt{2})M_{cr\,k}}{\sqrt{\frac{\omega_{cr\,k}}{\omega}} + \sqrt{\frac{\omega}{\omega_{cr\,k}}} + \sqrt{2}}. \tag{12}$$

The angular velocity, for which the component of torque due to the action of the k-th harmonic of Δ reaches a maximum, is equal to

$$\omega_{cr\,k} = \Delta_{ind}s_{cr\,k} = \frac{16(pk)^2\delta_e^2}{D^4\gamma}\frac{\mu}{\mu_0}. \tag{13}$$

The maximum torque component due to the action of the k-th harmonic component of Δ is equal to

$$M_{cr\,k} = 10.5\times10^{-6}\frac{D^3 l}{8kp}\Delta^2_{max}A_e. \tag{14}$$

Expressions (13) and (14) show that $M_{cr\,k}$, due to the action of the highest harmonics of Δ, even for a linear load $\Delta_{max\,k}$, equal to $\Delta_{max\,1}$, is $\underline{k}$ times smaller than the torque due to the action of the first harmonic component. Slip for which $M_{cr\,k}$ is reached, is k^3 times greater than $s_{cr\,1}$. Usually, $s_{cr\,3}$, $s_{cr\,5}, \ldots$ lie far beyond the limits of the values of slip which occur in the clutch under operating conditions. Because of this, in the range of operating speeds, the torque is mainly due to the action of the first harmonic component of Δ.

The mechanical characteristics can be obtained by summing the ordinates of the mechanical characteristics $M_k = f(\omega)$ which are plotted for each harmonic of Δ (Figs. 4 and 5), $M_s = \sum_{k=1,3\ldots} M_{ks}$. Some divergence between the calculated and experimentally determined characteristics (Fig. 4) is explained by the fact that in the derivation, variation of the relative magnetic permeability of the rotor with ω_{el} was not taken into account.

It is a characteristic of this type of clutch that the maximum in its mechanical characteristic is attained at considerably higher relative velocities, and is strongly smoothed out due to the action of the highest harmonics of Δ (Figs. 4 and 5). In accordance with (4) and (11), the critical torque, which is determined by the fundamental harmonic of Δ, is equal to

$$M_{\mathrm{cr}\,1} = 17 \times 10^{-9} \frac{DlA_{\rho}}{\delta} (Iw_{\mathrm{p}})^2 \frac{\sin^2\left(\frac{pa}{D}\right)}{p\left(\frac{a}{D}\right)^2} = Cf\left(p, \frac{a}{D}\right).$$

In this way, the smaller the ratio $\frac{a}{D}$ is, the larger is $M_{\mathrm{cr}\,1}$, and, in the limit,

$$M_{\mathrm{cr}\,1} \to 17 \times 10^{-9} \frac{Dl}{\delta} p\,(Iw_{\mathrm{p}})^2 A_{\rho} \text{ while } \frac{a}{D} \to 0.$$

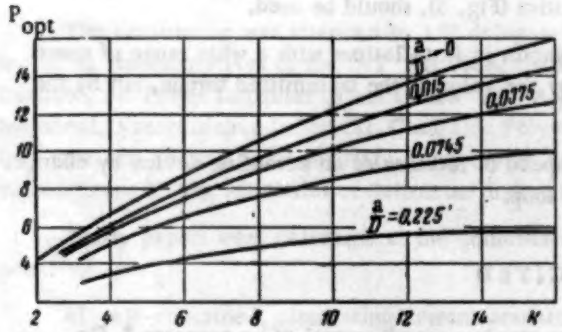


Fig. 6. Curves of the relationship $p_{\mathrm{opt}} = f(k)$ with $a/D = \mathrm{const}$.

To increase the torque, the spacing between poles (or intersecting teeth) must be minimized. The minimum value of a is limited by the increase of the leakage flux between the poles. Usually one selects $a = (10-20)\,\delta$. This condition is always met in clutches with finger-type inductors, since in these the change in induction in the gap occurs over a very small portion of the length of the inductor. The optimum number of pole-pairs p_{opt} can be determined from the condition for maximum torque by equating to zero the derivative, with respect to p, of the expression

$$M = 56.3 \times 10^{-9} \frac{DlA_{\rho}Iw_{\mathrm{p}}^2}{\delta} \frac{\sin^2\left(\frac{ap}{D}\right)}{\left(\frac{a}{D}\right)^2\left(\frac{K}{\frac{1}{p^2}} + \frac{p^2}{K} + \sqrt{2}p\right)} = C'f'\left(\frac{a}{D}, K\right).$$

The optimum number of pole-pairs depends, in this case, not only on the magnitude of K, determined by Equation (7), but also on the ratio $\frac{a}{D}$ (Fig. 6).

The curves $p_{\mathrm{opt}} = f(k)$, for several values of $\frac{a}{D}$, were plotted to determine p_{opt} (Fig. 6).

Example. Let us determine the p_{opt} for an AC of the PMS-0.4 type, with a toothed inductor.

Given: $D = 14.4$ cm, $\delta = 0.03$ cm, $a = 0.3$ cm, $n_{\mathrm{ind}} = 1340$ rpm, $s_n = 0.1$, $\gamma_{80^\circ C} = 0.312 \cdot 10^5\ \mathrm{ohm}^{-1}\mathrm{cm}^{-1}$, $\mu = 10^3$, $\mu_0 = 4\pi \cdot 10^{-9}$ h/cm. The calculations are carried out in the following order:

$$\omega_{\mathrm{ind}}s = 2\omega_{\mathrm{ind}}s_n = 25.6\ 1/\mathrm{sec}^{-1} \quad K = D^2/4\delta_1\sqrt{\gamma\mu_0/\mu\omega_{\mathrm{ind}}s} = 5{,}4,$$
$$a/D = 0{,}3/14{,}4 = 0.0208.$$

From the curve $p_{\mathrm{opt}} = f(k)$ (Fig. 6) corresponding to $\frac{a}{D} = 0.0208$, for the value $K = 5.4$, we find $p_{\mathrm{opt}} = 6-7$.

SUMMARY

The maximum torque in an AC with a phaneropolar inductor with intersecting teeth is greater than in clutches with finger-type and nonphaneropolar inductors.

It is necessary to select the smallest ratio of $\frac{a}{D}$ to increase the torque in an AC.

The mechanical characteristics of nonphaneropolar clutches have very sharply defined maxima.

The sharp mechanical characteristics of an AC with a phaneropolar and toothed (with intersecting poles) inductor increase its efficiency of transmitting the nominal torque. A clutch of this type is most useful for applications where large values of transmitted torque are required.

In clutches with diameters $D \leq (10 - 20)$ cm, the distribution of the required ampere-turns of excitation, for a phaneropolar and toothed inductor with intersecting poles, leads to complications and often makes it necessary to increase the clutch dimensions. In this case, a clutch of very simple design, with a finger-type inductor, possessing relatively smooth mechanical characteristics (Fig. 5), should be used.

It is also expedient to use an AC with a finger-type inductor in installations with a wide range of speed variations, where the clutch dimensions are determined not by the value of the transmitted torque, but by the necessary area of the heat-emitting surface.

When used in systems for the automatic control of the speed of rotation of an actuating device by changes in the driving current, the AC must contain a speed feedback loop.

LITERATURE CITED

[1] Ia. B. Rozman, "Electromagnetic slip clutch for the stepless speed control of an actuator," [In Russian] (Institute of Techno-economic Information, AN SSSR, 1954).

[2] E. M. Lopukhina, "Study of an asynchronous motor with a hollow, cylindrical rotor," Electrichestvo (USSR) 5 (1950).

[3] Alfred S. Cutman, "Designing an eddy current clutch servo," Mach. Design 10 (1950).

[4] L. R. Neiman, in the book: Surface Effect in Ferromagnetic Objects [In Russian] (Gosenergoizdat, 1949).

Received July 27, 1957

COORDINATION CONFERENCE ON AN AUTOMATED ELECTRICAL AC ACTUATOR WITH A SELF-CONTAINED SUPPLY

A coordination conference, organized by the Power Institute, A. N. Azerb. SSR and the Institute of Automation and Remote Control, AN SSSR, was held on October 7-9, 1957, in Baku, concerning an automatic, electrical, alternating current actuator with a self-contained supply.

The conference was attended by 120 delegates from 44 organizations in 14 towns in the Soviet Union, including representatives of the IAT AN SSSR, the Electromechanical Institute of AN SSSR, the Moscow Power Institute, the Power Institutes of AN Uzbek SSR, AN Latv. SSR, the L'vov Polytechnical, Azerbaidzhan Polytechnical, Azerbaidzhan Industrial, Cheliabin Polytechnical, Khar'kov Polytechnical, and Erevan Polytechnical Institutes, the Orden Lenina Military Academy of Tanks Corps, Giproneftmasha, Ts. K. B. "Electroprivod," Bashenergonefti, the Ministries of Petroleum Industries of Azerb. SSR, etc.

Thirty papers were delivered at the conference. The proceedings of the conference were divided into two parts:

a) self-contained, alternating current actuators;

b) automatic regulation and control of alternating current actuators.

In the introductory speech, the vice-president of A. N. Azerb SSR, Professor Z. I. Khalilov noted the significance of the conference, and formulated the problems which have to be solved.

In the paper of Academician V. S. Kulebakin, "Problems of automatic, combined regulation of electrical actuators," the theory and methods of designing systems with combined regulation by means of reaction to deviations and changes, were elucidated. It was shown that by using the combined method of control and regulation, separate coordinates of an automatically-operating system can be made almost independent of external actions or loads.

The theory and practice of throttle and frequency control systems was the subject of the second paper of Academician V. S. Kulebakin, "On the application of saturable reactors and rectifiers in widely controlled actuator systems." The speaker discussed the main circuits of direct current actuators, supplied from single-phase and three-phase power supplies, and controlled by means of saturable reactors, static characteristics of electric actuators, as well as circuits with saturable reactors and rectifiers for the frequency control of alternating current motors.

The results of scientific studies of self-contained, alternating current actuators carried on at the Power Institute of A. N. Azerb. SSR were presented by the Corresponding Member of A. N. Azerb. SSR, A. A. Efendizade.

The Institute has carried out studies of the automation of drill-feed on petroleum installations using rotary boring, and the automation of the brake machine control in drilling oil bores. A study was made of the operation of an asynchronous motor when supplied by a synchronous generator of commensurate power, as well as a system for the regulation of the excitation of synchronous generators in self-contained installations.

The paper of T. Z. Portnoi (Ts. K. B. "Elektroprivod") was devoted to self-contained electrical drives for machines in drilling installations. He presented the results of industrial tests of a direct current, diezel-electrical drive unit, compared direct current and alternating current diesel -electric actuators, and formulated the problems in designing diesel-electric drives for drilling installations.

The report of V. N. Bogoiavlenskii (IAT AN SSSR) "Regulated asynchronous actuator, using planetary transmissions, in self-contained supply systems" was devoted to a frequency control system, incorporating electromechanical transmission of alternating current, developed for an electric tractor.

The report of V. G. Vovka (Construction Bureau VNII) was also devoted to the application of electromechanical transmission.

Supply systems with frequency control by means of commutating generators was the subject of the papers read by representatives of IEM AN SSSR.

The report of V. P. Andreev and V. A. Prozorov, "An automated, alternating current actuator with very smooth starting and stopping," treated the problems of automating the control of an electric actuator with large flywheel loads, by means of an alternating current motor-generator system using a compensated commutator generator.

The reports of V. V. Rudakov "The work of IEM AN SSSR on quick-response, electromechanical frequency converters," and E. M. Smirnov "A comparison analysis of some electromechanical systems for the excitation of a compensated, commutator generator," examined various electromechanical frequency converters with an inverted or normal commutator, supplied from the windings of electrical machines or static transformers.

In the report of A. A. Dartau, "On the selection of an optimum system for the excitation of a commutator generator," experiments carried out on resonant systems for the excitation of commutator generators were described. An optimum type of commutator generator with resonance switching of field coils, was recommended.

Iu. M. Aleksandrov, in his paper "Single-phase, regulated, alternating current actuator with frequency control," described investigations of an ac motor-generator set suitable for mechanisms not subjected to sharp load surges, and consisting of a single-phase commutator generator, based on an amplidyne, and an asynchronous motor connected single-phase.

The report of V. M. Mamedov, "The operation of a longitudinal field, dynamoelectric amplifier with ac self-excitation," set forth the results of work directed at the design of alternating current amplifiers based on standard types of direct current, dynamoelectric amplifiers.

S. V. Strakhov (MEI) delivered a report on the calculation of the static stability of diesel- electrical, alternating current actuators. A method is given for the determination of the static stability in the absence, as well as in the presence of voltage and current feedback loops. It is shown that the limit of the static stability for an asynchronous motor supplied by a generator of commensurate power-rating, shifts appreciably in the direction of increasing slip, in comparison to the case of a motor fed by an infinite power network.

E. A. Iakubaitis (Power and Electrical Engineering Institute A. N. Latv. SSR), in his report, "Self-contained power station for a railway passenger coach," presented results of the operation of systems for the supply of electricity to passenger coaches, in the form of rectified ac, using a synchronous, contactless generator, and a system for the automatic control of the excitation of a self-exciting synchronous generator with variable rpm.

The problems of stabilizing the voltage of a synchronous generator, operating in a self-contained system, were treated in the reports of M. M. Rasulov (ENIN A. N. Azerb, SSR) "A system for the regulation of the excitation of synchronous generators in a diesel-electric actuator," V. A. Glebov (RII Zh. T.) "Forced starting of asynchronous motors in self-contained electrical systems," V. T. Zagorskii (L'vov Polytechnical Institute) "Capacitor start of short-circuited asynchronous motors from generators of commensurate power."

The report of M. M. Rasulov formulated the requirements, based on theoretical and experimental investigations, imposed on regulators of the excitation of synchronous generators in self-contained installations. The author presented the results of investigations of new circuits using saturable reactors for the control of the excitation of a generator.

V. A. Glebov's report presented the results of an investigation of the starting of a short-circuited asynchronous motor from a synchronous generator of commensurate power, for pulse starting, starting through an auto-transformer, and through a resistance.

V. T. Zagorskii described, in his report, the results of an analytical and experimental investigation of a synchronous generator — asynchronous motor system with series-connected, fixed capacitors. The speaker gave the conditions for the self-excitation and self-drive of the motor.

In the report of M. A. Nabiev (ENIN AN Azerb SSSR), "Investigation of transient processes in a synchronous generator — asynchronous motor system with regulated excitations," the theory of the control of a synchronous generator's excitation system was presented.

V. Iu. Nevraev (IAT AN SSSR) in his report "Analysis of transient processes in, and choice of structure for a system for the automatic voltage control of a self-contained synchronous generator — asynchronous motor installation," described a semi-graphical method of calculating the transient processes in an exciter — synchronous generator — asynchronous motor system.

V. A. Parail (Odessa Polytechnical Institute), in his report "An alternating current generator-motor system for diesel-electric actuators," presented a method for obtaining sharp regulation characteristics of an asynchronous motor, operating from a synchronous generator, by acting on the excitation circuit of the generator's exciter.

New types of supplies for boring installations were described by the representatives of Giproneftemasha, L. I. Shuturman in "Drive of drilling machines from mobile gas-turbogenerators," and I. N. Sulkhanishvili in "Diesel-electric, alternating current, drill drive, type DEB-2."

In these reports the authors outlined the aspects of designing a gas, turboelectric, drilling unit, pointed out the basic requirements imposed on diesel-electrical, drilling actuators, and presented the basic facts about an experimental drilling installation with gas-turboelectrical units, and about the circuit of a DEB-2 diesel-electrical actuator.

The report of I. I. Giul'mamedov (ENIN AN Azerb SSR) "Mechanical transient processes in an asynchronous motor supplied by a self-contained synchronous generator," contained the results of investigations of mechanical, transient modes in an asynchronous motor, in a self-contained system, under various load conditions, and the equations describing these modes were given.

B. A. Listengarten, in his report "The geometrical location of currents and voltages in an asynchronous motor operating from a synchronous generator of commensurate power," determined the location of currents and voltages in an asynchronous motor, taking into account the motor's magnetizing circuit, when working from a phaneropolar, constant excitation, synchronous generator.

In the second section, a report was given by Ia. B. Kadymov (ENIN AN Azerb. SSSR) "Stability criteria for automatic control systems, with distributed parameters," devoted to the popularization of frequency methods in the analysis of the stability of systems with distributed parameters. The report outlined an engineering method of investigating the stability of systems, described by complicated transcendento-algebraic equations (for example, systems with two delays).

Several reports were devoted to electrical actuators applicable to the petroleum industry.

K. N. Kulizade (Azerb. Industrial Institute), in his report "The application of low power, synchronous motors in installations of the oil industry," presented the results of the experimental application of a synchronous motor, with mixed excitation, and metallic rectifiers as a drive for pumping rigs on one of the oil fields in Azerbaidzhan.

The report of In. S. Kengerlinskii (ENIN AN Azerb. SSR) was devoted to the investigation of an automatic control system for tool-feed, as well as the results of investigations of a power regulating circuit, and a set of requirements, imposed on systems used in rotary drilling, were established.

A. M. Omarova (ENIN AN Azerb. SSR) described in her report the results of investigation and industrial trial of a brake machine (a synchronous generator operating in the region of dynamic braking) with a system for the automatic control of the lowering of an instrument (continuous and pulse).

The report of A. A. Mezhlumova (Azerb. Polytechnical Institute) "Electric drill — an electric drilling machine perfecting the technique of drilling," was devoted to the description of a machine which combined in one the electrical supply for the rotary table and the rotary table itself.

D. P. Petelin (IAT AN SSSR) delivered the report "Investigation of a system for the automatic control of the frequency and voltage of a synchronous generator" in which an investigation is made, based on the theory of multiple-connection control, of a system for the automatic control of the frequency and voltage of a synchronous generator, under the condition of any stable accuracy of each of the controlled quantities.

Arrangements for the oscillographic display of the angle θ in transient modes of synchronous machinery were described in the reports of G. M. Safarov (Azerb. Industrial Institute) and M. Ia. Ginzburg (ENIN AN Azerb. SSR).

Thirty-two persons took part in the discussion of the reports and the conclusions of the conference.

Several measures were planned in the conclusions of the conference, of which the following were the most important.

1. To intensify theoretical work on an automated electrical actuator with a self-contained supply, in scientific institutions of the Academy of Sciences, USSR, the Academies of Sciences of Soviet Republics, as well as in schools of higher education and branch scientific institutes.

2. To consider one of the basic problems in the field of self-contained diesel-electric drives to be the investigation of special versions of motors and generators which would ascertain the proper operation of actuators, and the investigation of complex systems for the control of the thermal and electrical parts of these actuators. The case of the simultaneous operation of several self-contained generating units with a common, sharply varying load, should also be considered.

3. To consider one of the most important theoretical and practical problems in the field of self-contained diesel-electrical actuators, to be the design of automatic systems for the remote control of prime movers; as well as the determination of optimum conditions and requirements for such control.

4. To devote proper attention, in the design of diesel-electric actuators, to the development of systems for the control of electric motors, making wide use of contactless automation, permitting the shaping of optimum load diagrams for the generator and prime mover.

5. In the investigation of transient processes in self-contained systems, it is necessary to take into account changes in the angular speed of the synchronous generator's rotor, and not to limit oneself to the development of semigraphical methods of investigation. Computers should be used widely.

Ia. B. Kadymov

ERRATA

In the note of M. A. Aizerman and F. R. Gantmakher, "On the stability of periodic modes in nonlinear systems with piecewise-linear characteristics," Automation and Remote Control 19, 6 (1958).

1. On p. 594 replace [1] everywhere by [4].

2. Number [4] in the Literature Cited should read:

[4] M. A. Aizerman and F. R. Gantmakher, "The stability of the linear approximation to the periodic solution of systems of differential equations with discontinuous right sides," Applied Mathematics and Mechanics (USSR) 21, 5 (1957).